

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2003, 13:31:44 ; Search time 3097 Seconds

(without alignments)
10233.437 Million cell updates/sec

Title: US-08-468-190c-1_COPY_9874_10962

Perfect score: 1089

Sequence: 1 ATGGCATCGCTACTTGAGAC.....CCGACCTCATATAAATCTGA 1089

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database :
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_iny:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rod:*
36: em_hlg_man:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlg_hum:*
40: em_hlg_mus:*
41: em_hlg_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1087.4	99.9	18912	14	ILU28832
2	63	5.8	1627	14	HSWGDLYCO
3	51.2	4.7	125020	9	AF429315
4	43.2	4.0	1246	6	AX164174
5	41.4	3.8	138025	8	AP003261
6	41.4	3.8	200183	8	AP003227
7	40.2	3.7	125020	9	AF429315
8	39.6	3.6	19277	2	DMU19909
9	39	3.6	110000	2	LMFICHR18_04
10	38.8	3.6	117004	2	AP003743
11	38.8	3.6	130732	2	AP003848
12	38.2	3.5	16984	1	AE006954
13	38.2	3.5	40051	1	MSGY224
14	38.2	3.5	40838	1	MTY25D10
15	38.2	3.5	109554	2	CNS08C90
16	38	3.5	4543	3	DROCSM
17	38	3.5	38889	3	DMBN25G24
18	38	3.5	45760	3	AC017610
19	38	3.5	177724	3	AC104146
20	38	3.5	291976	6	AX251969
21	37.6	3.5	6310	6	AX344361
22	37.6	3.5	6310	6	AX348754
23	37.6	3.4	1172	3	LE1P1BR1L
24	37.4	3.4	110000	2	LMFICHR32_20
25	37.4	3.4	110000	2	LMFICHR32_21
26	37.4	3.4	10177	1	AB004972
27	37.2	3.4	10177	1	AB004972
28	37	3.4	10261	1	AB011997
29	37	3.4	224800	1	SM5501789
30	37	3.4	346547	1	AP003012
31	36.8	3.4	30000	1	AB050904
32	36.6	3.4	6176	3	AF104400
33	36.6	3.4	12177	3	AF104401
34	36.6	3.4	71516	2	AC014220
35	36.6	3.4	155263	8	AP003224
36	36.6	3.4	160001	3	AP0113619
37	36.6	3.4	189976	3	AC023710
38	36.6	3.4	305018	3	AE003437
39	36.2	3.3	40476	1	SCC75A
40	36	3.3	10732	6	E32986
41	36	3.3	161205	2	AC106407
42	35.8	3.3	58270	9	AF250325
43	35.8	3.3	115666	2	AC105744
44	35.8	3.3	153472	9	AB020868
45	35.8	3.3	176967	2	AC022742

ALIGNMENTS

```
RESULT 1
LOCUS      ILU28832
DEFINITION Infectious laryngotracheitis virus us10, us2, protein kinase, UL47,
            glycoprotein G, ORF5, glycoprotein D, glycoprotein I, glycoprotein
            E, ORF9 genes, complete cds.
ACCESSION  U28832
VERSION    U28832.1 GI:1486484
KEYWORDS   .
SOURCE     Gallid herpesvirus 1
ORGANISM   Gallus gallus
REFERENCE  1 (bases 1 to 18912)
AUTHORS    Wild,M.A., Cook,S. and Cochran,M.
```

TITLE	A genomic map of infectious Jaryngotracheitis virus and the sequence and organization of genes present in the unique short and flanking regions
JOURNAL	Virus Genes 12 (2), 107-116 (1996)
MEDLINE	97033380
PUBMED	8879127
REFERENCE	2 (bases 1 to 18912)
AUTHORS	Wild, M.A.
TITLE	Direct Submission
JOURNAL	Submitted (09-JUN-1995) Martha A. Wild, Syntro Research Laboratory, 3535 General Atomics Court, San Diego, CA 92121, USA
FEATURES	Location/Qualifiers
source	1..18912
	/organism="Gallid herpesvirus 1"
	/strain="USDA challenge strain"
	/db_xref="taxon:10386"
repeat_region	<1..2909
	/note="large inverted repeat flanking the short unique region"
	/rpt_type="inverted"
CDS	697..1533
	/codon_start=1
	/product="US10"
	/protein_id="AAC55093.1"
	/db_xref="GI:1486485"
	/translation="MENMLDGCYPLALMDSPHITAHAVPCGERRRQGAAGVASESADS VDCIRIASRLRELSEISLKDGEFTSARDRNALIANERLSAAGSARAT GLGRLPMASVESVANSPTDNGNGELTEEMAGTEGDFWLSDGDRPDSRTM OSEMRVIEKEILSMLSRHLPALASERTSRSLAAGHMCCLMHPREKEDCID SIYVSLFCVGTGRVPOSEMRREYLAALRAGAANAASPEVSASIFADGIALALARR"
CDS	complement(1900..2784)
	/note="coding region unique to avian herpesviruses"
	/codon_start=1
	/product="unknown"
	/protein_id="AAC55094.1"
	/db_xref="GI:1486486"
	/translation="MSKCYCLARHLKSPRCVGRRAVFGGLATNSRPTSHLDAFSA AFRCTDLPGRFMRASQSCDIEFWPDLAAVVOAARAVFEGERLGSQVADITAD PRIAPAKRAVAALVAGLMTALSELVSGNGELSKVNGKOIPRAAMEIRDPVPI GPDLISFEESAVALPELVYIRAGGASHRSAMNNOSSAPAGLAIRIGMVFSLIY IALPLSNPLPEDLPBGSONSIRAFVAMNCVATDKIMSDVAVPVEBSYSICLAE IIMCRAPCYCPNPPPKM"
misc_feature	2910..16003
	/note="short unique region"
CDS	complement(2916..3605)
	/codon_start=1
	/product="US2"
	/protein_id="AAC55095.1"
	/db_xref="GI:1486487"
	/translation="MAPVKYIVASAVDSHKLPNSRELSDSGMKELVHAVKMASYD RPSYLSIVRNASLYEVSGELSLPRMCRPIYIRGEGGDPGVSPEMSGDAGYHLS SGAYAAKEFHLWLVGTADICMALNLPAPKTELTTERGKRGVEVFLVNGKRTL SLHSVMTTLAPSSLRTPMPYSIVKFLKXKPNAAVCVSDSDGEPKPTFLSLRK SKRPSRRRR"
CDS	3694..5124
	/codon_start=1
	/product="protein kinase"
	/protein_id="AAC55096.1"
	/db_xref="GI:1486488"
	/translation="MRRRIRCSRSRAKRRRTTENPLTSKRCVCLDSFSRMSLRPYA EILPTAGVEVRLAVSVLTERAEPTENTAVNSIPPANEGQNFAYAGDPSPTTK VDSHTDFDEASDYAGVPLAQTRLKHSDFLQHFVLPDLVAGAGTCDVGRYTE EEORRRGVNSTNOGSKCKRLIAKYVNGTRFASQLENELVIGRLNHEVLIQOELI RYPNTYMLTOROFOFLDSYVDEAFPMKDSFMLKOTRIMKOLMASVSYIHSKLIH RDLKLENIFLNCQKTYLGDGTYTPRENSREPREYMGCVTAANSPEITLARSYCI TDTRSCGVILEMYSHEFCPIGDGGGPHOOLKLVITSLSDCEDFPPOCNLNIYH YASIDRAGHTVPSLIRNLHLPADEYPLVMKLFDMWLRPSAAVLAMPLEFSAEERT ITIIGHKHKRIRREIRARVRSMSEG"
CDS	5210..7081
	/codon_start=1
	/product="UL47"
	/protein_id="AAC55097.1"
	/db_xref="GI:1486489"
	/translation="MTPLRLTKRPPARRCQSVFIHYSSEKLDRLRYNKKTLMPDST MRHSGGANHKGIRPRKRSIGAFSAREKGTGRNLTSSSSSDMLDPSDKEGKGM TVDDPADITAEVSOANDVLOLYVHEPAEERVYERKPPIOFNAMPQSPNADQ PTPRAPITPSSAEVVDARLKYGAFSVKYLQCVPTTRKSKSTVYDADQSLDSEWRH ONVTYVLRQHTLRLRHPSKSYVNCNDPLAVANOFHMGMRVPSLKASPEENI OHGMAAVFRNAGGLSEYFALVLAETVLAITMPHALVFMKALADGAFNDYDETRYID VODTSREFEPALGCLSEYFALVLAETVLAITMPHALVFMKALADGAFNDYDETRYID VKNSEYNGAGCTLRGIVASNTALAVCANVYSTIKRLEPVAVSACNVAVRETTLA RRPMSDRIYRIKOEFFEFYIAMORVANTHFCNITIKRSDCAPFLIPASSERKL OQLKMKCPILVPIQYEDFSKAMSELRKLETFYVAISDRPDRPSRLRLISDHR EITADVFRKRVDEPVASVALSTAAAGVKAKARLTISVRAVPAGAVASRRKSEI" 7245..8123
CDS	/codon_start=1
	/product="glycoprotein G"
	/protein_id="AAC55098.1"
	/db_xref="GI:1486490"
	/translation="MSGFSNIGSIATVSLVCSILCASVGLAPVLIDLESSPPEFGKI IAQACNRTTTEVTPWMSDYSQRTGEVSEVAKWFGNSNPESFVGVSESGSHEDLS TCNALINHLNLSVCRASDAGIPDQKCEKQVRLRSGVELGSYSGNSLVLYPGMY DAGIYAYQLSGYGGKGYGSLVDYGPBGCHDOYGYTYYSIADASDLSYDVASPEL DGMREEDYSNCDMPRLRPPTTVCSHVEDENATDLYILMDEACAPLDDEYDERSE TMRKRVFSPSTLQO" 8333..11290
CDS	/note="ORF5"
	/codon_start=1
	/protein_id="AAC55099.1"
	/db_xref="GI:1486491"
	/translation="MGTMVLRLFLAVADAAALPTGFCRYWKVPGCGTIOENLVIA ESPVGTGATYPPPEGAVSFOIADPTPLRLRYGATEDEBLEROTSADNDNFSIS YRPRBEIGHAVFTIGVFATGQSTESSYVSVSVNALSLEVSALTECDENFLQNP TWGSKRWLGGPSPYVRNDNVALTKAOLYIECYNSAAGCGLTSLNTPFEYSPIRYN VTTWTGGSPSRITVYSRENGOVLNNSGFLVYKTPDIDGRAMVIANVANSPADS GSVLAFTAPRREGKLPASAIOLHRIIDMSTEBPRTCTDCQKMETPYRALGNSPRD SIRGATLPPEDTAAPOPDGTSTPTTVPRATTILIPRSTSMGFFSTARAGST LSVPOETDRLTSTPLPLPTEPESNTLPTAPISLETPSAHEHTQOAEVTV LVQSPSTESERTLASQOEPMYFTQSTSOALTOIQALETALFTQTPASQMETT OPGAETEAPOAPSTIPEITQSGRSTPPETASAPAEVLTQSTTVEVFPQTS TVPTLTSSSTSEPIAIFRTIOSAGEATQTESASAPDPROSTETHEFTQVSPVPA TORPSTEGVDSPTPEPRTFLTGAPEIOTQTPAAEAVYRSSSTMETRQST LASONPTSSGCTHNTBPRYPNOTTQTKLYTENKTLSPPTVSRFHMTAESQT PLIDVATVEKFSNDGEVYATCVSTYASPTREVENKVDLYVDWDEISGSPACVFS NEKQKOLYTRVDGRTSVOLMCISYSHSPEPCLDTSLIAEKDIABEPLFTSDP OTAYCTTLPSGVVPRFWSLNVSLPEYLTATVTSRTAGOSTVMKSSARAGAWIS GRGNIIECTVLISDGTRTYTRKERCLTNWIAVENGAQAOLYSLRSGVSLCGSI SALYATLMTAIFY" 11098..12402
CDS	/codon_start=1
	/product="glycoprotein D"
	/protein_id="AAC55100.1"
	/db_xref="GI:1486492"
	/translation="MHRPHLRHSRYAAKEVNLKHMDCGKRCGGAAYTLEMTVC RIMEHLCFVRNMDRHLFLRNAMWTIVLISFASOSTAATVYDYILGRALALATIP AVGRYNNYLTVRSKGCYVELNPLISANDDMISAKEKEKGPPEASVYVYIKGDG EDKYCPILYKREYRECGVQLSEBCAVOAMAVDYPSTLVSRNGGLTFISPTAAL SGQYLTLTKLIGRAQTALVTLVENDRLKIGSONLFPSCWTEGDOYQTFQGEHLV IADTNTRHADVYGYEDILORMNNILRKNKNSPSPADPVSQPEIAPVAKKAGRRP DAESEKKAPPEDESDMOAEGENPAALPEDEVEDPEHDDPNSDPDYNDMPAY IYVEETKSSNAVSMPIFAAFVACAVALLVGLWVSIYKARS" 12510..13598
CDS	/codon_start=1
	/product="glycoprotein I"
	/protein_id="AAC55101.1"
	/db_xref="GI:1486493"
	/translation="MSLSLGLTLLAATLAPGAMGIVITGNHVSARIDDDHIVYAP RPEATIDQLFPMGCPHPRPGSYAVARSDITNOCYBELSEERENTHSSSVF VGVKVEYITFSASRLNGLPHPKRLTIRNRPNDGSEFYIVRLDDKKEITDVAIDL SVQOFANTATRGRLYSKASCRTGLPVLQLEALVLTREESRNQAVATVATTSAPAL TPTPTVATIGSELAEHFTFPMLELNDVDYDEPTPASTENSVTVRLGTMSPTLIGTV AVAYSATIGLIVISIVTRNMCTPHRLDLVQDDDEERSQTVRESRKFGMVACEIK GAPODSELVELVAIVNSALSPDSIKM" 13792..15291
CDS	/codon_start=1

gene	/db_xref="taxon:10386"	
	/note="also known as infectious laryngotracheitis virus (ILTIV)"	
TATA_signal	144..1517	/gene="gd"
	144..151	/gene="gd"
CDS	326..1441	/gene="gd"
	1507..1517	/codon_start=1
	/gene="gd"	/product="glycoprotein D"
		/protein_id="AA98925.1"
		/db_xref="GI:493596"
		/translation="MDRLFLRNAPFMTVLLSPASOSTAAVTVDTILGRALDALTPAVPANNRYLTVRSRGCEIVEINLSINVDIMSAKKEKGPPEASVSPYVTKGGDGEDKPCYIKREYRECCGDVQLSCAQSQAOMALDVSSTLVSRNGAGLTIESPNARLSGGYLLFLTKIGRFAQTALVTVLEVNDKLTGTSKPNLPSPICWTEIYQTFGEHELTPADATNTRHADVDYRREDIQLRNWNLILRKNPAPRPDSVQELPAVTKAEGRLYPDASSEKKAPEDESDMQAEASGEPMALPEDEVEDETHEDPMSDDPYNDMPAPVPEETKSSNAVSMPIFAFVACAAVSMQVWGAS"
polyA_signal	1507..1517	/gene="gd"
BASE COUNT	410 a 428 c 428 g 361 t	
ORIGIN		
Query Match	5.8%; Score 63;	DB 14; Length 1627;
Best Local Similarity	100.0%; Pred. No. 1.7e-06;	
Matches 63; Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	ATGGCATTCCGCTACTTGGAACTCTGGCTCTCTCTTGGCCGGAGCGTCCGACCCCTTGGCGCG 60
Db	1565	ATGGCATTCCGCTACTTGGAACTCTGGCTCTCTCTTGGCCGGAGCGTCCGACCCCTTGGCGCG 1624
Qy	61	ATG 63
Db	1625	ATG 1627
RESULT 3		
AF429315	125020 bp	DNA linear PRI 18-JAN-2002
LOCUS	AF429315	
DEFINITION	Homo sapiens junctophilin 3 (JPH3) gene, partial cds.	
VERSION	AF429315	
KEYWORDS	AF429315.1 GI:17646244	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE	1 (bases 1 to 125020)	
AUTHORS	Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S., Ingwersoll,Ashworth,R.G., Flesher,A., Stevanin,G., Brice,A., Potter,N.T., Ross,C.A. and Margolis,R.L.	
TITLE	A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2	
JOURNAL	Nat. Genet. 29 (4), 377-378 (2001)	
MEDLINE	21583737	
PUBMED	11694876	
REFERENCE	2 (bases 1 to 125020)	
AUTHORS	Holmes,S.E., Ingwersoll,Ashworth,R.G., Ross,C.A. and Margolis,R.L.	
TITLE	Direct Submission	
JOURNAL	Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA	
FEATURES	Location/Qualifiers	
source	1..125020	/organism="Homo sapiens"
	/db_xref="taxon:9606"	
	/chromosome="16"	
	/map="16q24.3; between D16S520 and WI-12410"	
	/note="Isolated from a patient with Huntington's disease-Like 2 (HDL2)"	
	complement(35581..35746)	
	/rpt_type=tandem	

[illegible]

RESULT 4
AX164174/c 1246 bp DNA linear PAT 22-JUN-2001
LOCUS AX164174 Sequence 4 from Patent WO0138564.
DEFINITION AX164174
ACCESSION AX164174
VERSION AX164174.1 GI:14545112
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1246)
AUTHORS Rouleau, G.A., Latreille, R.G., Rochefort, D., Cossette, P. and Ragsdale, D.
TITLE Local for idiopathic generalized epilepsy, mutations thereof and method using same to assess, diagnose, prognosis or treat epilepsy
JOURNAL Patent: WO 0138564-A 4 31-MAY-2001;
McGill University (CA)
FEATURES
source location/Qualifiers
1. 1246
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 112 a 36 c 104 g 109 t 885 others
ORIGIN
Query Match 4.0%; Score 43.2; DB 6; Length 1246;
Best local similarity 12.7%; Pred. No. 0.63;
Matches 39; Conservative 141; Mismatches 127; Indels 0; Gaps 0;
QY 65 GAATCGGATACCTGGAATGACGCTCCGCCAGATTGACGATGACATGATGATG 124
DB 445 SHKSBSSSYWYKAAACNNHATMBHCBABSCBSCBNCBASPDMCMTPTKXSY 386
QY 125 TCGGCCCTCCGCCGAGAGATGATGATGATGATGATGATGATGATGATGATG 184
DB 385 SASNSNTCDYBHYSHNASHDHHTNHSCHMYTCSSAIVSYSSCSDSSSYWYRAY 326
QY 185 CCCCAAAACCTACTCAGAGACCGCTCCGCCGCTGCTGATATACCAACGAT 244
DB 325 NCSYCMYTHSSMHMCCSMYMYNYMYSMSMSSTSSSHYCTYSDBTATATTT 266
QY 245 GCTACCAAGAACTTAGGAGAGAGCGCTTTGAATTCATCATCATCATCATCAT 304
DB 265 MMKTYATNRTKTBTRNSCBRRACCTTATNMMAKISTWASHARCNMYCTMBGRK 206
QY 305 TTGCGGCTGTAAGAGACGAGTACGCTTCCGCCGAGACAGACTACCGACTC 364
DB 205 GCTHSSNCGTTCRDYSHRSMHMBANACNRNANNSDSITNMGNTNCKTBSGAKHS 146
QY 365 CACACCC 371
DB 145 MMBSTCB 139
RESULT 5
AP003261 138025 bp DNA linear PLN 10-AUG-2002
LOCUS AP003261 Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
DEFINITION PAC clone: P0471B04.
ACCESSION AP003261
VERSION AP003261.3 GI:22202658
KEYWORDS
SOURCE Oryza sativa (japonica cultivar-group) (cultivar: Nipponbare) DNA,
clone: P0471B04.
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriophytidae; Oryzae; Oryza.
REFERENCE 1
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Oryza sativa Nipponbare (G33) genomic DNA, chromosome 1, PAC
clone: P0471B04

JOURNAL
REFERENCE 2 (bases 1 to 138025)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasakienias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
On Aug 9, 2002 this sequence version replaced gi:21202837.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI Nonredundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI Nonredundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown protein'. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from SP6 to T7 of the PAC clone.
This sequence of P0471B04 clone has an overlap with P0698A10 clone
(DBJ: AP003287) at the position 1 to 24,438 of 5' end and an
overlap with P0018C10 clone (DBJ: AP003227) at the position 48,801
to 138,025 of 3' end. Detailed information on overlap and assembly
quality together with annotation of this entry is available at
http://rgp.dna.affrc.go.jp/Genomeseg.html.
Location/Qualifiers
1. 138025
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="P0471B04"
/join(9533..9715,9887..10019,10850..10988,11365..11401,
13640..13931,14823..14955,15564..15576)
/gene="P0471B04.1"
/gene(9533..9715,9887..10019,10850..10988,11365..11401,
13640..13931,14823..14955,15564..15576)
/gene="P0471B04.1"
/note="hypothetical protein
similar to Arabidopsis thaliana chromosome 1, F10C21_14"
/codon_start=1
/protein_id="BAC07317.1"
/db_xref="GI:22202659"
/translation="MSPPEVIGQFGDTTYRKFVGLAMETQETMRKRYEFGEGELLE
AVITDKNTGRSKGYGFTRPDPAMACVDAPVIDGRANCMALAGVGRSRPT
POHGARSFRVMSFSQOAGIGGAGAPSPHATPFAVAPGGLYHYGVSPYSDY
SYPTNYNYTGAOYDFGAAAMAAAGVWSSPFYFPOFGSGSTTTNAGSG
GYNLOYPOMHPSTVASTAAVTFPAOYGGPLSLAASPQAOAAKRGKSNPNMAL
TGGCFMLIANNMRTKNSKOLIEQCHFEISRP"
/complement(join(17199..17342,17914..18008,18141..18178,
18784..19073))
/gene="P0471B04.2"
/complement(join(17199..17342,17914..18008,18141..18178,
18784..19073))
/gene="P0471B04.2"
/codon_start=1
/product="putative SnRK1-interacting protein 1"
/protein_id="BAC07318.1"
/db_xref="GI:22202860"
/translation="MMSGALAKLLGRRAAATPTLAPPAAKARARRRPPAPSPED
EDDFAGGVAAAPATGISTPLAEVRLGRGKRVENLYTRVDPESLKYIPMHIVRK
IINHAPWSEGVRSIVYSSDGSVYIRITLNGTAEIYRATGTSADDTGCGP
YOKAAEMAFRACRIGLGLHLEHMS
join(24593..24925,25029..25403)
/gene="P0471B04.3"

[illegible][illegible]

[illegible][illegible]

Combining Monsanto and RGP-Japan sequencing data.

NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES

source

1.130732"Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="7"

/clone="OJ1715-A07"

BASE COUNT 36776 a 28066 c 27679 g 38160 t 51 others

ALIGN

Query Match

Best Local Similarity 56.2%; Score 38.8; DB 2; Length 130732;

Matches 73; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY

622 GCCAGGAGGCGACGACGACGCGCGGCGACACCCGCGCTCACTCAAC 681

Db 122232 GCCAGGCGGCGGAGCGGACGAGGAGGACACCCGCGCGCTCACTTGGC 122173

QY

682 AGCCCTCCGACCTTAACGACACTTTTCTTCCCTGCTAGAAAATGCGTGGAT 741

Db 122172 AACCTCTGATCTGCATCTGCATCTGAGCGCTGCTCTCTCAACCTCGCGAGGTG 122113

QY

742 CATTCGACAC 751

Db 122112 GAGATCGATC 122103

RESULT 12

AE006954/c

LOCUS AE006954 16984 bp DNA linear BCT 27-APR-2001

DEFINITION Mycobacterium tuberculosis CDC1551, section 40 of 280 of the

ACCESSION AE006954 AE000516

VERSION AE006954.1 GI:13880047

KEYWORDS

SOURCE

ORGANISM

Mycobacterium tuberculosis CDC1551.

Mycobacterium tuberculosis CDC1551.

Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Corynebacterineae; Mycobacteriaceae;

Mycobacterium; Mycobacterium tuberculosis complex.

1 (bases 1 to 16984)

Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O.,

Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E.,

Kolony, J.F., Nelson, M.C., Umayam, L.A., Ermolaeva, M.,

Salzberg, S.L., Delcher, A., Uterback, T., Weidman, J., Khouri, H.,

Gill, J., Mikula, A. and Bishai, W.

Whole genome comparison of Mycobacterium tuberculosis clinical and

laboratory strains

Unpublished

2 (bases 1 to 16984)

Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O.,

Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E.,

Kolony, J.F., Nelson, M.C., Umayam, L.A., Ermolaeva, M.,

Salzberg, S.L., Delcher, A., Uterback, T., Weidman, J., Khouri, H.,

Gill, J., Mikula, A. and Bishai, W.

Direct Submission

Submitted (25-APR-2001) The Institute for Genomic Research, 9712

Medical Center Dr., Rockville, MD 20850, USA

Location/Qualifiers

FEATURES

SOURCE

/organism="Mycobacterium tuberculosis CDC1551"

/strain="CDC1551"

/db_xref="taxon:83331"

/note="clinical strain"

174. 1406

/gene="MF0538"

174. 1406

/gene="MF0538"

/note="similar to GB:D31821.SP:P21542; identified by

sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="acyltransferase, putative"

/protein_id="AAK44761.1"

/db_xref="GI:13880048"

/transl_table=11

/translation="MRAETGTEIRALGLRIYAANVVFHFRPMLGDSGCFDAL

APVLDGAGQVDLFTLSGFTVLTWNTLDRKGRSWSRANLFLMLRLATWVTVTL

HLAAWVIFTLHGVSPSPAGQLTAISVQIILVOLWQPTFDSGSDGSPASISA

EMLAVLLEGLILVIFRPMKHAIRAGLMLAFASLPPVLLASGQFTYPSWLPRI

VYOFAGALACAAVRRLRPTDRRIAGYSVLGVAIVGILYLHAHPLACVEDSG

VVDVLPVLYSLAIGVSLPALSTRLMVGGOISFCLYVHVELVHTMGVAVOYE

LALDOPFKMNVVGLAIALGAAILLYHFVEEPGRMRMRMDVKAAARSEPEPVG

STRQIDDALEGVSARAV"

1538. 2233

/gene="MF0539"

1538. 2233

/note="identified by Glimmer2; putative"

/codon_start=1

/transl_table=11

/product="hypothetical protein"

/protein_id="AAK44762.1"

/db_xref="GI:13880049"

/translation="MSRPGTYIGLTLVGLVGVNPGCPSTYRPLTLDYRLNPVAVIG

DSYTTGDEGLIGSGKSWTAQTOMLAAGVRIADVAEGRAGYGVPGHGGVFDLP

RAVQPDALVPEPSSNDGQDPPDPLAEKVRDIFDLARRRAPSALVIAPAMP

TADVPQMLRTRDYLGQARAAGNVFVPIDHMFVDRPELICADGVHNDGHEYLA

DKIAPLSMKELVQ"

complement(2285. 2401)

/gene="MF0540"

complement(2285. 2401)

/note="identified by Glimmer2; putative"

/codon_start=1

/transl_table=11

/product="hypothetical protein"

/protein_id="AAK44763.1"

/db_xref="GI:13880050"

/translation="MSPMLAFRTAAAPLAGILLOLARLERWYERHAED"

complement(2522. 3424)

/gene="MF0541"

complement(2522. 3424)

/note="identified by Glimmer2; putative"

/codon_start=1

/transl_table=11

/product="hypothetical protein"

/protein_id="AAK44764.1"

/db_xref="GI:13880051"

/translation="MLRRCAGTBRGIRMPMDLRRALLRMGAGAGAGVAVAP

GALVDPLEPOAPAPPEPPAGSSLSLPTSSPISARAGIKTNVYISMPPOSGOLR

PVYALHGKDNAGKMLDLGVEGLARLYEKRPFAVAVGAGTNYWHRRSGGSGA

WYDELPLMTSGMDTSRGFLGWSMGYCALLLGRALPAPATAGICISPALETSP

TGSTPGAFSDYDYOHSVYIGLPALNSIPLAVDGTSDRFRYFRQFVNLQHPAGS

SPGCHASVWRDLPGELAAAS"

3605. 3955

/gene="MF0542"

3605. 3955

/gene="MF0542"

/note="similar to GP:2792343; identified by sequence

similarity; putative"

/codon_start=1

/transl_table=11

/product="methyltransferase-related protein"

/protein_id="AAK44765.1"

gene /db_xref="GI:13880052"
/translation="MGCSITCINLSEVPNENKRNQALDLSIRVIRHSPDIE
PSSGIDVWSDALIHADPRKRVLEEARVLRPGELLFTDPMDADVDPGVLDPYID
RINLRDLSMRPYA"
3948..4253
/gene="MT0543"
3948..4253
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44766.1"
/db_xref="GI:13880053"
/translation="KREAQALGFVEVLDQRLVNRNLRTHYSRVEEELRRLIEGKS
SDEYLDKMKVGLKNVNEADNGSHVGHPTPEKRLPICOLPTPAIDSTAGRRRYR"
4388..5692
/gene="MT0544"
4388..5692
/note="similar to SP:P46349 PID:969027 PID:1708639
GG:AL009126; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="amino acid permease"
/protein_id="AAK44767.1"
/db_xref="GI:13880054"
/translation="MIAIGVIGAGLFEVSGVIRATGPAFLTYALCALIYLVMRM
LGEAANDSTGAFADYAAKALGWAFGSVMLYFWIVYGFVEAAGGVLTYTID
ALMLTSLCGLMMNTATNLVSSSGEEFEPWAGKATVIGFLVGLTAFKGLLPFH
GMDNFSLHSGEFPFDGAVGAFVATVYAFEMTQEVTTAAARPPORQANST
VYARIVPFGSVFLIYILPRNSLELQASPYVALRMKIGGDQIMNAVYLRATVS
CTNGLGTSRRLFYLARKOEAPQVLVNRGVPTRAIMGSSVVGFLCYIMAWSPSA
TVFELNNSGALVILEYLLIALSQILVNRQTSQNLGVRMMLRPGSLIYTVGIVAV
LARKAFDPAARQQLSLISMAVVGCTLVTTVLRRLNRPW"
complement(5676..6080)
/gene="MT0545"
complement(5676..6080)
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44768.1"
/db_xref="GI:13880055"
/translation="WVGQQLPQWLARENRVTNFIQRLMAGLPAFLIEHVGRRSK
PRTLRNLFSSADVRAGVAILITYGPRDLKIKTTAAGGRMRRYCTGVA NPRRL
TKAEAPYVSSRWRRVFAFLPDEAVLLTKAD"
6185..7573
/gene="MT0546"
6185..7573
/note="similar to GB:U00018 SP:P46716 PID:467041;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="glutamate-1-semialdehyde 2,1-aminomutase"
/protein_id="AAK44769.1"
/db_xref="GI:13880056"
/translation="MGSEFQATSVRGAARTSAQLEFACSVIRPGVSPVRAFTANG
GTPRRTTAHGCMLIDADGKNRYVDLVCWKGMITGHHAPVAYEVAARAKRGISFGAP
TPALEQLAGSEITIGRVAPEIRILVNSGTEATWASVRLARGETGAKIVKSGCYHGV
DALDOLDSGVALTGLCDPQRPASPRSOSRGILSSPGVGAADOTIVLPYNDIDA
VOQTARGEGDIAAVITETASPGNMGVPPGCFPAALRAITAEHAGLLIDDEVWGR
VRSQWQSIDVPPADLFAFGKVMGSGMPAAAFGGRAEVMORLAPGVYQAGTLGNDP
VAVAGLATTLEAADDAVYATLADANDRLAGLSALTDAYVPHOISRAGMLSYFGE
TPYVDFASARKSQTRIPAFPHAMDAGVITPPCSAFEMFVSALDDAARGLANALP
AAARAAQGERA"
7573..8181
/gene="MT0547"
7573..8181
/note="identified by match to protein family HMM"

Query Match	Best Local Similarity	Score	DB 1	Length	DB 2	Score	DB 3	Length	DB 4	Score	DB 5	Length	DB 6	Score	DB 7	Length	DB 8	Score	DB 9	Length	DB 10	Score	DB 11	Length	DB 12	Score	DB 13	Length	DB 14	Score	DB 15	Length	DB 16	Score	DB 17	Length	DB 18	Score	DB 19	Length	DB 20	Score	DB 21	Length	DB 22	Score	DB 23	Length	DB 24	Score	DB 25	Length	DB 26	Score	DB 27	Length	DB 28	Score	DB 29	Length	DB 30	Score	DB 31	Length	DB 32	Score	DB 33	Length	DB 34	Score	DB 35	Length	DB 36	Score	DB 37	Length	DB 38	Score	DB 39	Length	DB 40	Score	DB 41	Length	DB 42	Score	DB 43	Length	DB 44	Score	DB 45	Length	DB 46	Score	DB 47	Length	DB 48	Score	DB 49	Length	DB 50	Score	DB 51	Length	DB 52	Score	DB 53	Length	DB 54	Score	DB 55	Length	DB 56	Score	DB 57	Length	DB 58	Score	DB 59	Length	DB 60	Score	DB 61	Length	DB 62	Score	DB 63	Length	DB 64	Score	DB 65	Length	DB 66	Score	DB 67	Length	DB 68	Score	DB 69	Length	DB 70	Score	DB 71	Length	DB 72	Score	DB 73	Length	DB 74	Score	DB 75	Length	DB 76	Score	DB 77	Length	DB 78	Score	DB 79	Length	DB 80	Score	DB 81	Length	DB 82	Score	DB 83	Length	DB 84	Score	DB 85	Length	DB 86	Score	DB 87	Length	DB 88	Score	DB 89	Length	DB 90	Score	DB 91	Length	DB 92	Score	DB 93	Length	DB 94	Score	DB 95	Length	DB 96	Score	DB 97	Length	DB 98	Score	DB 99	Length	DB 100	Score	DB 101	Length	DB 102	Score	DB 103	Length	DB 104	Score	DB 105	Length	DB 106	Score	DB 107	Length	DB 108	Score	DB 109	Length	DB 110	Score	DB 111	Length	DB 112	Score	DB 113	Length	DB 114	Score	DB 115	Length	DB 116	Score	DB 117	Length	DB 118	Score	DB 119	Length	DB 120	Score	DB 121	Length	DB 122	Score	DB 123	Length	DB 124	Score	DB 125	Length	DB 126	Score	DB 127	Length	DB 128	Score	DB 129	Length	DB 130	Score	DB 131	Length	DB 132	Score	DB 133	Length	DB 134	Score	DB 135	Length	DB 136	Score	DB 137	Length	DB 138	Score	DB 139	Length	DB 140	Score	DB 141	Length	DB 142	Score	DB 143	Length	DB 144	Score	DB 145	Length	DB 146	Score	DB 147	Length	DB 148	Score	DB 149	Length	DB 150	Score	DB 151	Length	DB 152	Score	DB 153	Length	DB 154	Score	DB 155	Length	DB 156	Score	DB 157	Length	DB 158	Score	DB 159	Length	DB 160	Score	DB 161	Length	DB 162	Score	DB 163	Length	DB 164	Score	DB 165	Length	DB 166	Score	DB 167	Length	DB 168	Score	DB 169	Length	DB 170	Score	DB 171	Length	DB 172	Score	DB 173	Length	DB 174	Score	DB 175	Length	DB 176	Score	DB 177	Length	DB 178	Score	DB 179	Length	DB 180	Score	DB 181	Length	DB 182	Score	DB 183	Length	DB 184	Score	DB 185	Length	DB 186	Score	DB 187	Length	DB 188	Score	DB 189	Length	DB 190	Score	DB 191	Length	DB 192	Score	DB 193	Length	DB 194	Score	DB 195	Length	DB 196	Score	DB 197	Length	DB 198	Score	DB 199	Length	DB 200	Score	DB 201	Length	DB 202	Score	DB 203	Length	DB 204	Score	DB 205	Length	DB 206	Score	DB 207	Length	DB 208	Score	DB 209	Length	DB 210	Score	DB 211	Length	DB 212	Score	DB 213	Length	DB 214	Score	DB 215	Length	DB 216	Score	DB 217	Length	DB 218	Score	DB 219	Length	DB 220	Score	DB 221	Length	DB 222	Score	DB 223	Length	DB 224	Score	DB 225	Length	DB 226	Score	DB 227	Length	DB 228	Score	DB 229	Length	DB 230	Score	DB 231	Length	DB 232
-------------	-----------------------	-------	------	--------	------	-------	------	--------	------	-------	------	--------	------	-------	------	--------	------	-------	------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------

RESULT 14
MTY25D10/C 40838 bp DNA linear BCT 03-AUG-2001
LOCUS
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 28/162.
ACCESSION 295558 AL123456
VERSION 295558.1 GI:3261781
KEYWORDS
SOURCE
ORGANISM Mycobacterium tuberculosis H37Rv.
Mycobacterium tuberculosis H37Rv.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacteriaceae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 40838)
Cote, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eigmeier, K., Gas, S., Barry III, C.E., Tekale, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Fellwell, T., Gentles, S., Hamlin, N., Holtroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, S., Squares, S., Squires, R., Sulston, J.E., Taylor, K., Whitehead, S. and Barrett, B.G.
TITLE
Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence
JOURNAL
Nature 393 (6685), 537-544 (1998)
MEDLINE
98295987
PUBMED
9634230
REFERENCE
2 (bases 1 to 40838)
Parkhill, J.
TITLE
Direct Submission
SUBMITTED (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
COMMENT
On Jun 27, 1998 this sequence version replaced gi:2113983.
NOTES:
Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web
(URL: http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have been renumbered from the original cosmid submissions but the old gene designations are in brackets after the new gene numbers.
Gene prediction was based on a hidden Markov Model of 18 genes implemented in TParse (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an upstream ribosome (atg, gta, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.
Location/Qualifiers
1. 40838
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
<1. 480
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
/clone="Y20G10"
1. 40095
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
/clone="Y25D10"
complement(312..707)
/gene="RV0523c"
complement(312..707)
/gene="RV0523c"
/note="RV0523c, (MTCY25D10.02, unknown), len: 131 aa,

gene
CDS
misc_feature
gene
CDS
gene
CDS
some similarity to M. tuberculosis MTCY336.06 006592 (136 aa); fasta scores, opt: 197 z-score: 236.8 E(): 5e-08, (38.4% identity in 99 aa overlap)"
/transl_table=11
/product="hypothetical protein RV0523c"
/protein_id="CA08990.1"
/db_xref="GI:2113985"
/db_xref="SPTREMBL:O06389"
/translation="MQLPQAPRVRVYVTPVQRIQVAGLPAFALIEHVGSRKPYR
TRPQVSAVDGRAGVAILITVGRNRMINKITIAAGGMRRTGTFVAPNRLITTA
EATVYSRRRPFYATLPDEAVILITKAD"
821..2209
/gene="hcmL"
821..2209
/note="RV0524, (MTCY25D10.03), len: 462, hcmL, highly similar to GSA MYCIE P46716 glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8) (GSA) (446 aa), fasta scores, opt: 1532 z-score: 1581.4 E(): 0, (82.6% identity in 460 aa overlap); contains P500600 Amino transferases class-III pyridoxal-phosphate attachment site"
/codon_start=1
/transl_table=11
/product="hcmL"
/protein_id="CA08991.1"
/db_xref="GI:2113986"
/db_xref="SWISS-PROT:O06390"
/translation="MGTEQATSRVRCARTRSAQLFEACSVIPGVNSPVRAFTAVG
GTPREITEAHGCMILIDAGNRVYDVCSPMILIGHAVPAVEAVAKAAGLSEFAP
TPAETOLAGIIGRAVAPERIRLVNSGTEAVSLAGFTGRKIVKFGSCVGHV
DALIADAGSGVATIGLCDPQPPASPSQSRGIPSPVTAADVITVPIYNDIDA
VQGFAPRFGQIAVYTESAGNGVYPPGQGNALRITAEHALITLDEVTGFR
VSRGSGWIDPDPADLPFGVMSGGGAPAFGRVQWRLVAPGYQGTGSCGP
VAVAGATITLRADADAVYTAIDANADLALGIEALDVLVPHQISRAGNNLSVFEE
TPYDPSARASQTRVYPAFPAHMLDAGVYPPCSAEAFVSAALDDAAPERIANALP
AAARAAQPERPA"
1616..1726
/gene="hemL"
/note="P500600 Amino transferases class-III pyridoxal phosphate attachment site"
2209..2817
/gene="RV0525"
2209..2817
/note="RV0525, (MTCY25D10.04), len: 202 aa, unknown, highly similar to M. leprae Q49821 B2168_C3-276 (202 aa) opt: 1151 z-score: 1349.4 E(): 0, (82.5% identity in 200 aa overlap), also similar to SLD0395 SYNECHOYSTIS SP. Q55734 hypothetical 23.8 kDa protein (212 aa) opt: 207 z-score: 251.8 E(): 5.1e-07 (28.2% identity in 195 aa overlap), also some similarity to Y019_MYCTV_Q10512 hypothetical 39.2 kD protein cy427.09 (364 aa) opt: 236 z-score: 282.0 E(): 1.1e-08, (34.3% identity in 198 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein RV0525"
/protein_id="CA08992.1"
/db_xref="GI:2113987"
/db_xref="SPTREMBL:O06391"
/translation="MPEQOVYVVRHGSVNPVPGIIGRLPFGHLSATGCAQAAVAD
ALADRIDVAVIASPIQAOETAPLAARHDLAVETPDIDIESANFEERRVGPDCGAW
RDPVWMOALNPPTSPKCEPVYDIAARHTAYVAKAVRAGVEVVCVSHQLPVITLFL
YTGCKRLMDPRRDCALASTSLYGDRLVDVYVYSGNAL"
2832..3482
/gene="RV0526"
2832..3482
/gene="RV0526"
/note="RV0526, (MTCY25D10.05), len: 216, unknown, highly similar to M. leprae Q49816 U218C (216 aa) opt: 1144 z-score: 1259.4 E(): 0, (78.5% identity in 214 aa overlap), some similarity to Thiodoxins eg RESA_BAC5U

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2003, 13:03:03 ; Search time 324 Seconds
(without alignments)
7569.217 Million cell updates/sec

Title: US-08-468-190c-1_COPY_9874_10962

Perfect score: 1089
Sequence: 1 ATGCGATCGCTACTGTGAC.....CCGACTCAATATAAATGTGA 1089

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseg-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1089	100.0	13473	16	AAT33504	Infectious laryngo
2	1089	100.0	13473	17	AAT44384	Infectious laryngo
3	1087.4	99.9	18912	17	AAT44385	Infectious laryngo
4	346	31.8	534	16	AAT33513	Infectious laryngo
5	38.2	3.5	4403765	22	AA199683	Mycobacterium tube
6	38.2	3.5	4411529	22	AA199682	Mycobacterium tube
7	38	3.5	2526	22	AA26455	D.melanogaster cor
8	38	3.5	2677	23	AB106927	Drosophila melanog
9	38	3.5	2838	23	AB18521	Drosophila melanog

10	38	3.5	4528	23	ABL21135	Drosophila melanog
11	38	3.5	10587	23	ABL18520	Drosophila melanog
12	38	3.5	19468	23	ABL06926	Drosophila melanog
13	38	3.5	21399	23	ABL21134	Drosophila melanog
14	37.6	3.5	813	24	ABO13968	Oligonucleotide fo
15	37.6	3.5	813	24	ABO13969	Oligonucleotide fo
16	37.6	3.5	6310	24	ABL70322	Human gene regulat
17	37.6	3.5	6310	24	AA561269	Chemically treated
18	37.6	3.5	6310	24	ABK13365	Signal transductio
19	36.6	3.4	6151	23	AB106707	Drosophila melanog
20	36.6	3.4	9173	23	AB106706	Drosophila melanog
21	36.2	3.3	5059	23	AA84332	Streath virus nucl
22	36	3.3	10732	21	AA10594	Gene encoding a su
23	35.6	3.3	904	18	AA16484	Myceliorhiza the
24	35.2	3.2	640	24	ABO37730	Oligonucleotide fo
25	35.2	3.2	640	24	ABO37731	Oligonucleotide fo
26	35.2	3.2	927	22	AAH90774	CPE 78 coding sequ
27	35.2	3.2	951	22	AAH90883	2CFE 78 coding seq
28	35.2	3.2	2814	21	AA250960	A. halophila beta1
29	35.2	3.2	19718	19	AAV52322	Streptococcus pneu
30	35	3.2	864	23	AA593326	DNA encoding novel
31	35	3.2	2257	23	AA593627	DNA encoding novel
32	35	3.2	2546	23	AA590013	EST clone EL360.
33	34.6	3.2	475	20	AAV88421	Human secretory po
34	34.6	3.2	50341	19	AAV22674	DNA sequence of a
35	34.6	3.2	50341	21	AAZ39519	LS shuttle phasid
36	34.6	3.2	52297	16	AAZ39519	Myobacteriophage
37	34.6	3.2	52297	16	AAZ39519	LS myobacteriophag
38	34.6	3.2	52298	14	AAO47357	Human DNA for a no
39	34	3.1	2855	22	AA531515	Human polynucleoti
40	34	3.1	2855	22	ABO66839	Lolium perenne LPO
41	33.4	3.1	775	24	ABN87171	DNA encoding novel
42	33.4	3.1	2895	24	AA599365	DNA encoding novel
43	33.4	3.0	2895	24	AA599337	Human ATPase-ssoc
44	33.2	3.0	498	24	ABN76781	Drosophila melanog
45	33.2	3.0	601	23	ABL09003	

ALIGNMENTS

RESULT 1	AAAT33504	standard; DNA: 13473 BP.
ID	AAAT33504	standard; DNA: 13473 BP.
XX	AAAT33504;	
XX	18-NOV-1996	(first entry)
XX	Infectious laryngotracheitis virus short region sequence.	
XX	Infectious laryngotracheitis virus; ILTV; herpesvirus;	
XX	attenuation; vector; vaccine; chicken; poultry; immunisation; ds.	
XX	Infectious laryngotracheitis virus.	
XX	Key	Location/Qualifiers
XX	repeat_region	1..273
XX	msc.feature	/tag- a
XX		/function- internal repeat region
XX		274..13371
XX		/tag- b
XX		/function- unique short sequence of ILTV
XX		complement (281..970)
XX		/tag- c
XX		/label- US2-gene
XX		1059..2489
XX		/tag- d
XX		/label- Protein-kinase_gene
XX		2575..4107
XX		/tag- e
XX		/label- UL47-1like_gene
XX		4113..4445
XX	CDS	

```
FT      /*tag- f.  
FT      /label- ORF4  
FT      complement (4139..4519)  
FT      /*tag- g  
FT      /label- ORF4_reverse-complement  
FT      misc_difference 4535  
FT      /*tag- h  
FT      /note- "base 4535 is given as 's' in the  
FT      specification"  
FT      CDS  
FT      4609..5487  
FT      /*tag- i  
FT      /label- gG_gene  
FT      5697..8654  
FT      /*tag- j  
FT      /label- g60_gene  
FT      CDS  
FT      complement (6948..7826)  
FT      /*tag- k  
FT      /label- ORF6_reverse-complement  
FT      8462..9766  
FT      /*tag- l  
FT      /label- gD_gene  
FT      9874..10962  
FT      /*tag- m  
FT      /label- gI_gene  
FT      CDS  
FT      complement (10617..11150)  
FT      /*tag- n  
FT      /label- ORF8_reverse-complement  
FT      11159..12658  
FT      /*tag- o  
FT      /label- gE_gene  
FT      12655..13447  
FT      /*tag- p  
FT      /label- ORF10  
FT      CDS  
FT      13002..13003  
FT      /*tag- q  
FT      /note- "bases 13002-13303 are given as 'ss' in  
FT      the specification"  
FT      repeat_region  
FT      13372..13473  
FT      /*tag- r  
FT      /function- terminal repeat region  
FT      W09508622-A1.  
FT      30-MAR-1995.  
FT      16-SEP-1994: 94WO-US10628.  
FT      24-SEP-1993: 93US-0126597.  
FT      (SYTR ) SYNTRO CORP.  
FT      Cochran MD, Wild MA:  
FT      WPI: 1995-139591/18.  
FT      P-PSDS: AAM00630, AAM00631, AAM00632, AAM00633, AAM00634, AAM00635,  
FT      AAM00636, AAM00637, AAM00638, AAM00639, AAM00640, AAM00641 W00642.  
FT      Recombinant attenuated infectious laryngotracheitis virus - for use  
FT      in vaccines to protect poultry from infection from the virus, also  
FT      methods of distinguishing between vaccinated and naturally infected  
FT      blids  
FT      Example 1: Page 79-94; 177p; English.  
FT      The unique short region (AAT33504) of infectious laryngotracheitis  
FT      virus (ILTV) genomic DNA contains genes (see also AAT33505 and  
FT      AAT33510-13) that are associated with ILTV virulence. A deletion in  
FT      those genes, esp. the glycoprotein g6 gene, glycoprotein g1 gene,  
FT      thymidine kinase gene, us2 gene, uI47-like gene or the glycoprotein  
FT      g60 gene, will attenuate the ILTV. A gene for a foreign antigen may  
FT      be inserted into the us2, uI47-like, ORF4, g6, g60 or g1 gene to  
FT      produce a recombinant ILTV. Recombinant ILTV may be used as a  
FT      multivalent vaccine, esp. for use in poultry.
```

```
XX      SQ      Sequence 13473 BP; 3390 A; 3582 C; 3547 G; 2951 T; 3 other;  
SQ      Query Match      100.0%; Score 1089; DB 16; Length 13473;  
SQ      Best Local Similarity 100.0%; Pred. No. 0;  
SQ      Matches 1089; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY      1      ATGCATCGCTACTGGACCTGCTCTCTCTCCGCGAGCGCTCGCACCTTCGGCGCG 60  
DB      9874      ATGCATCGCTACTGGACCTGCTCTCTCTCTCCGCGAGCGCTCGCACCTTCGGCGCG 9933  
OY      61      ATGGGAATCGTGATCACTGGAATACAGTCTCCGCCAGATTGACGATCAATCATGTG 120  
DB      9934      ATGGGAATCGTGATCACTGGAATACAGTCTCCGCCAGATTGACGATCAATCATGTG 9993  
OY      121      ATGCTCGCGCTCGCCCGGAGAGCTACAAATTCACGTATTTTTCATGCTGCGCAG 180  
DB      9994      ATGCTCGCGCTCGCCCGGAGAGCTACAAATTCACGTATTTTTCATGCTGCGCAG 10053  
OY      181      AGACCCGACAAACCTACTAGAGAACCGTCCGCGCTGCTGCTGATATTAACAAC 240  
DB      10054      AGACCCGACAAACCTACTAGAGAACCGTCCGCGCTGCTGCTGATATTAACAAC 10113  
OY      241      CAGTGTACAGAGAACTTACGAGAGAGCGCTTTGAAAATTGCATCATGCTCTTCT 300  
DB      10114      CAGTGTACAGAGAACTTACGAGAGAGCGCTTTGAAAATTGCATCATGCTCTTCT 10173  
OY      301      GTTTTGTGCGCTGCTAAAGTACCGGATACAGTCTCCGCTCGGAGACGATCAACCGGA 360  
DB      10174      GTTTTGTGCGCTGCTAAAGTACCGGATACAGTCTCCGCTCGGAGACGATCAACCGGA 10233  
OY      361      CCTCCACACCCGTTTAAAGTCACTATACGAATCTCGTCCGAGACGACGCGGATGTC 420  
DB      10234      CCTCCACACCCGTTTAAAGTCACTATACGAATCTCGTCCGAGACGACGCGGATGTC 10293  
OY      421      TACGTAATTTGTTGCGCTAGACAGACACCAAGAACCCATTACGCTTTCGGATCAACTA 480  
DB      10294      TACGTAATTTGTTGCGCTAGACAGACACCAAGAACCCATTACGCTTTCGGATCAACTA 10353  
OY      481      TCGGTATATCAATTCGCGAAGACCGCGCGCTCGCGAGCTATTCGAAAGGCTTCGTGT 540  
DB      10354      TCGGTATATCAATTCGCGAAGACCGCGCGCTCGCGAGCTATTCGAAAGGCTTCGTGT 10413  
OY      541      CGCACCTTGGATTAATCTACCGTCAACTTGAGGCTATCTCAGAGCCGAGAAAGTTGG 600  
DB      10414      CGCACCTTGGATTAATCTACCGTCAACTTGAGGCTATCTCAGAGCCGAGAAAGTTGG 10473  
OY      601      CGCACTGCGCAAGGCTACGTTGCCACGAGGAGCCACGACGACGAGCGCGAGGACAAAC 660  
DB      10474      CGCACTGCGCAAGGCTACGTTGCCACGAGGAGCCACGACGACGAGCGCGAGGACAAAC 10533  
OY      661      CCGAGCGCGCTACCTGCAACGAGCGCTCGCAACTTGAAACGGAACCTTTACCTTCCC 720  
DB      10534      CCGAGCGCGCTACCTGCAACGAGCGCTCGCAACTTGAAACGGAACCTTTACCTTCCC 10593  
OY      721      TGGCTAGAAAATGGCGTGATCATTTACGAACCGACCGCAACGAAAATTTCAAACTTT 780  
DB      10594      TGGCTAGAAAATGGCGTGATCATTTACGAACCGACCGCAACGAAAATTTCAAACTTT 10653  
OY      781      ACTGTGCGTTCGCGGACAAATGAGCCCTACCTAATTTGGGTAAACCGGTGCGCGCTG 840  
DB      10654      ACTGTGCGTTCGCGGACAAATGAGCCCTACCTAATTTGGGTAAACCGGTGCGCGCTG 10713  
OY      841      AGCGCAAGATCGGCGCTGCTATTTGTAATTTTCATGCTACCAAGAAAATGTGCACCCG 900  
DB      10714      AGCGCAAGATCGGCGCTGCTATTTGTAATTTTCATGCTACCAAGAAAATGTGCACCCG 10773  
OY      901      CACGGAATAATTAGACAGGCTCTGCGACAGACGAGAAAGCTTCCAACTAGAAAGGAA 960  
DB      10774      CACGGAATAATTAGACAGGCTCTGCGACAGACGAGAAAGCTTCCAACTAGAAAGGAA 10833  
OY      961      TCGGGAATAATTGAGCCATGTTGCTGCGGAATTAACAAGGCGCTGACCAAGATAGT 1020  
DB      10833      TCGGGAATAATTGAGCCATGTTGCTGCGGAATTAACAAGGCGCTGACCAAGATAGT 10896
```



```

Db      10834 TCGGGAAATTTGGACCCAGTGTGCTGCGAAATTAACAGGCGCTGACCAAGATAGT 10893
OY      1021 GAACCTTGAGACCTGTTGGCATTTGTTAACCCGCTTGGCCTTAAGCTCCGCCAGACTCAATA 10880
Db      10894 GAACCTTGAGACCTGTTGGCATTTGTTAACCCGCTTGGCCTTAAGCTCCGCCAGACTCAATA 10953
OY      1081 AAAATGTGA 1089
Db      10954 AAAATGTGA 10962

RESULT 2
AAAT4384
ID      AAAT4384 standard; DNA; 13473 BP.
XX
XX      AAAT4384;
XX
XX      02-JUN-1997 (first entry)
XX
XX      Infectious laryngotracheitis virus unique short region.
XX
XX      ILTV; vaccine; vector; attenuation; poultry;
XX      avian infectious bronchitis virus; Newcastle disease virus;
XX      infectious bursal disease virus of chickens;
XX      Marek's disease virus; herpesvirus; ss.
XX
XX      Infectious laryngotracheitis virus USDA strain 8302.
XX
XX      Key
XX      Location/Qualifiers
XX      repeat_region
XX      1..273
XX      /tag= a
XX      274..13371
XX      /tag= b
XX      /note= "unique short region"
XX      151..166
XX      /tag= c
XX      /note= "polyA signal for ORF1"
XX      complement (281..970)
XX      /tag= d
XX      /label= ORF1(RC)
XX      /note= "unique-like short 2 (US2) gene"
XX      1007..1010
XX      /tag= e
XX      /note= "TATA signal for ORF1 and ORF2"
XX      1040..1043
XX      /tag= f
XX      /note= "TATA signal for ORF1"
XX      1042..1045
XX      /tag= g
XX      /note= "TATA signal for ORF2"
XX      1059..2489
XX      /tag= h
XX      /label= ORF2
XX      /note= "protein kinase gene"
XX      2491..2496
XX      /tag= i
XX      /note= "polyA signal for ORF2"
XX      2538..2543
XX      /tag= j
XX      /note= "polyA signal for ORF2"
XX      2575..4107
XX      /tag= k
XX      /label= ORF3
XX      /note= "unique long 47 (UL47)-like gene"
XX      4523..4526
XX      /tag= l
XX      /note= "TATA signal for ORF4"
XX      4113..4445
XX      /tag= m
XX      /label= ORF4
XX      complement (4139..4519)
XX      /tag= n
XX      /label= ORF4(RC)
XX
XX      TATA_signal
XX      4523..4526
XX      /tag= o
XX      /note= "TATA signal for ORF5"
XX      4609..5487
XX      /tag= p
XX      /label= ORF5
XX      /note= "glycoprotein gg gene"
XX      4609..4686
XX      /tag= q
XX      4687..5484
XX      /tag= r
XX      5564..5569
XX      /tag= s
XX      /note= "polyA site for ORF5"
XX      5653..5658
XX      /tag= t
XX      /note= "polyA signal for ORF5"
XX      5697..8654
XX      /tag= u
XX      /label= ORF6
XX      /note= "glycoprotein g60 gene"
XX      6987..7727
XX      /tag= v
XX      /note= "repeat region consists of approx. 23
XX      repeats of 30-36 bp"
XX      complement (6948..7826)
XX      /tag= w
XX      /label= ORF6(RC)
XX      8455..8458
XX      /tag= x
XX      /note= "potential TATA signal for ORF7"
XX      8461..9766
XX      /tag= y
XX      /label= ORF7
XX      /note= "glycoprotein gd gene, alternative start
XX      codon at 8633..8635"
XX      9819..9822
XX      /tag= z
XX      /note= "TATA signal for ORF8"
XX      9874..10962
XX      /tag= aa
XX      /label= ORF8
XX      /note= "glycoprotein gi gene"
XX      9874..9939
XX      /tag= ab
XX      9940..10959
XX      /tag= ac
XX      complement (10617..11150)
XX      /tag= ad
XX      /label= ORF8(RC)
XX      11069..11072
XX      /tag= ae
XX      /note= "TATA signal for ORF9"
XX      11159..12658
XX      /tag= af
XX      /label= ORF9
XX      /note= "glycoprotein ge gene"
XX      11159..11212
XX      /tag= ag
XX      11213..12655
XX      /tag= ah
XX      12483..12486
XX      /tag= ai
XX      /note= "TATA signal for ORF10"
XX      12865..13447
XX      /tag= aj
XX      /label= ORF10
XX      13572..13743
XX      /tag= ak
XX
XX      CDS
XX      26-SEP-1996.

```



```

FT      repeat_region 1..2909
FT      /tag- a
FT      /note- "short repeat region"
FT      CDS
FT      697..1533
FT      /tag- b
FT      /label- SRORF2
FT      /note- "US10 gene"
FT      CDS
FT      complement (2916..3605)
FT      /tag- c
FT      /label- SRORF1
FT      /note- "short repeat open reading frame 1"
FT      misc_RNA
FT      2910..16003
FT      /tag- d
FT      /note- "unique short region"
FT      polyA_signal
FT      2796..2891
FT      /tag- e
FT      /note- "polyA signal for ORF1"
FT      CDS
FT      complement (2916..3605)
FT      /tag- f
FT      /label- ORF1(RC)
FT      /note- "unique-like short 2 (US2) gene"
FT      TATA_signal
FT      3642..3645
FT      /tag- g
FT      /note- "TATA signal for ORF1 and ORF2"
FT      TATA_signal
FT      3675..3678
FT      /tag- h
FT      /note- "TATA signal for ORF1"
FT      TATA_signal
FT      3677..3680
FT      /tag- i
FT      /note- "TATA signal for ORF2"
FT      CDS
FT      3694..5124
FT      /tag- j
FT      /label- ORF2
FT      /note- "protein kinase gene"
FT      polyA_signal
FT      5126..5131
FT      /tag- k
FT      /note- "polyA signal for ORF2"
FT      polyA_signal
FT      5173..5178
FT      /tag- l
FT      /note- "polyA signal for ORF2"
FT      CDS
FT      5210..7081
FT      /tag- m
FT      /label- ORF3
FT      /note- "unique long 47 (UL47)-like gene"
FT      TATA_signal
FT      7158..7161
FT      /tag- n
FT      /note- "TATA signal for ORF4"
FT      CDS
FT      6748..7080
FT      /tag- o
FT      /label- ORF4
FT      /note- complement (6774..7154)
FT      CDS
FT      complement (6774..7154)
FT      /tag- p
FT      /label- ORF4 (RC)
FT      TATA_signal
FT      7158..7161
FT      /tag- q
FT      /note- "TATA signal for ORF5"
FT      CDS
FT      7245..8123
FT      /tag- r
FT      /label- ORF5
FT      /note- "glycoprotein gg gene"
FT      sig_peptide
FT      7245..7332
FT      /tag- s
FT      /note- "mat_peptide"
FT      mat_peptide
FT      7333..8120
FT      /tag- t
FT      /note- "polyA_signal"
FT      polyA_signal
FT      8200..8205
FT      /tag- u
FT      /note- "polyA site for ORF5"
FT      polyA_signal
FT      8289..8294
FT      /tag- v
FT      /note- "polyA signal for ORF5"
FT      CDS
FT      8333..11290
FT      /tag- w
FT      /label- ORF6

```

```

FT      repeat_region
FT      /note- "glycoprotein 960 gene"
FT      9623..10363
FT      /tag- x
FT      /note- "repeat region consists of approx. 23
FT      repeats of 30-36 bp"
FT      CDS
FT      complement (9584..10462)
FT      /tag- y
FT      /label- ORF6(RC)
FT      TATA_signal
FT      11091..11094
FT      /tag- z
FT      /note- "potential TATA signal for ORF7"
FT      CDS
FT      11098..12402
FT      /tag- aa
FT      /label- ORF7
FT      12453..12448
FT      /tag- ab
FT      /note- "TATA signal for ORF8"
FT      CDS
FT      12510..13598
FT      /tag- ac
FT      /label- ORF8
FT      /note- "glycoprotein gi gene"
FT      sig_peptide
FT      12510..12575
FT      /tag- ad
FT      /note- "mat_peptide"
FT      mat_peptide
FT      12576..13595
FT      /tag- ae
FT      /note- complement (13253..13786)
FT      CDS
FT      /tag- af
FT      /label- ORF8(RC)
FT      TATA_signal
FT      13705..13708
FT      /tag- ag
FT      /note- "TATA signal for ORF9"
FT      CDS
FT      13792..15291
FT      /tag- ah
FT      /label- ORF9
FT      /note- "glycoprotein ge gene"
FT      sig_peptide
FT      13792..13845
FT      /tag- ai
FT      /note- "mat_peptide"
FT      mat_peptide
FT      13846..15288
FT      /tag- aj
FT      /label- ORF9
FT      TATA_signal
FT      15116..15119
FT      /tag- ak
FT      /note- "TATA signal for ORF10"
FT      CDS
FT      15288..16080
FT      /tag- al
FT      /label- ORF10
FT      repeat_region
FT      16004..18912
FT      /tag- am
FT      CDS
FT      16129..17013
FT      /tag- an
FT      /note- complement (17380..18216)
FT      CDS
FT      /tag- ao

```

W09629396-A1.
 26-SEP-1996.
 21-MAR-1996.
 06-JUN-1995.
 23-MAR-1995.
 (SYTR) SYNPRO CORP.
 Cochran MD, Wild MA.
 WPI: 1996-443372/44.
 P-FSDB: AAM06782, AAM06783, AAM06784, AAM06785, AAM06786, AAM06787,
 AAM06788, AAM06789, AAM06790, AAM06791, AAM06792, AAM06793, AAM01415,
 AAM01416.
 Recombinant infectious laryngotracheitis virus with deletion in the
 glycoprotein G, gi or US2 gene, etc. - useful for vaccines against
 infectious laryngotracheitis in poultry

Example 11; Page 138-154; 216pp; English.

CC The nucleotide sequence of 19,912 bp of contiguous DNA (AAAT44385)
CC from the unique short and flanking region of infectious
CC lagovirhaecicilis virus (ILTV) contains the entire 13,098 bp unique
CC short region (see also AAAT44384) and includes 17 open reading frames
CC that encode proteins (AAW06782-94, AAW01415-16) of over 100 amino
CC acids, 10 of which show homology to other virus genes. Novel
CC recombinant, attenuated ILTV comprises the ILTV genome cont. a
CC deletion in the unique short region, esp. in the glycoprotein g5,
CC gi. US2, ORF4, UL47-like or g60 gene. The attenuated virus is
CC useful as a vaccine against ILTV. A foreign gene encoding an
CC antigen e.g. from another avian virus can be inserted into the US2,
CC UL47-like, ORF4, g5, g60 or gi gene to provide a multivalent
CC vaccine for chickens and other poultry. Deletion of the g5 or gi
CC gene provides as a negative marker to distinguish vaccinated from
CC infected animals.

Sequence 18912 BP; 4416 A; 5261 C; 5251 G; 3984 T; 0 other;

Query Match	99.9%	Score 1087.4	DB 17	Length 18912
Best Local Similarity	99.9%	Pred. NO. 0		
Matches 1088	0	Mismatches	1	Indels 0
				Gaps 0

QY	1	ATGGCAATCGCTACTTGGAACTCGTGGCTCCCTTGGCGGACAGCTGCACCTTTCGGCGG	60
Db	12510	ATGGCAATCGCTACTTGGAACTCGTGGCTCCCTTGGCGGACAGCTGCACCTTTCGGCGG	125689
QY	61	ATGGGAATCGTGATTCACCTGGAAATCAACGTCCTCCGCGAGATTGACAGCATCATCTGTG	120
Db	12570	ATGGGAATCGTGATTCACCTGGAAATCAACGTCCTCCGCGAGATTGACAGCATCATCTGTG	126239
QY	121	ATGTGCGCGCTCGTCCGCCGGAAGCTACAAATTCAATCTGACCTATTTTCTATGGCTGCGCG	180
Db	12630	ATGTGCGCGCTCGTCCGCCGGAAGCTACAAATTCAATCTGACCTATTTTCTATGGCTGCGCG	126899
QY	181	AGACCCCAAAACCCCTACTCAGGAACGTCGCGCTGCGCTTTCGCTGATTAACAAC	240
Db	12690	AGACCCCAAAACCCCTACTCAGGAACGTCGCGCTGCGCTTTCGCTGATTAACAAC	12749
QY	241	CAGTGGCTACCGAGAACTGTACGAGGAGCGCTTGAATAATTGCAATCATGATGATCGCTTCT	300
Db	12750	CAGTGGCTACCGAGAACTGTACGAGGAGCGCTTGAATAATTGCAATCATGATGATCGCTTCT	12809
QY	301	GTTTTTGTGCGCTGTAAAGTGACCGAGTACACGTTCTCCGCTCGAACAAGATTAACCGGA	360
Db	12810	GTTTTTGTGCGCTGTAAAGTGACCGAGTACACGTTCTCCGCTCGAACAAGATTAACCGGA	12869
QY	361	CGTCCACACCCGTTTAAAGTCACATATACGAATTCCTCGTCCGAAAGACACGCGGATGTTC	420
Db	12870	CGTCCACACCCGTTTAAAGTCACATATACGAATTCCTCGTCCGAAAGACACGCGGATGTTC	12929
QY	421	TACGTAATTGTGGGCTAGACGACACCAAGAACCCATTGAGCTCTCCGATTCACACTA	480
Db	12930	TACGTAATTGTGGGCTAGACGACACCAAGAACCCATTGAGCTCTCCGATTCACACTA	12989
QY	481	TCGGTGTATCAATTCGCGAACACCCGCGGACTCGCGACTCTATTCACAGGCTTCGTCT	540
Db	12990	TCGGTGTATCAATTCGCGAACACCCGCGGACTCGCGACTCTATTCACAGGCTTCGTCT	13049
QY	541	CGACACTCGGATTACTCTACCGTCACAACTTGAGGCGCTATCTCAGGACGAGGAAGTTGG	600
Db	13050	CGACACTCGGATTACTCTACCGTCACAACTTGAGGCGCTATCTCAGGACGAGGAAGTTGG	13109
QY	601	CGCAACTGGCAAGCTGACGTTGGCAAGGAGGCGACAGACAGCAGCGCGAGGCGCAAC	660
Db	13110	CGCAACTGGCAAGCTGACGTTGGCAAGGAGGCGACAGACAGCAGCGCGAGGCGCAAC	13169
QY	661	CGAGCGCCGCTCATCTGCAACCAAGCGCTCCGAATCTGAAGCGGAACACTTACCTTCCC	720
Db	13170	CGAGCGCCGCTCATCTGCAACCAAGCGCTCCGAATCTGAAGCGGAACACTTACCTTCCC	13229

Oy	721	GGCGTACAAAATGGCCGTGATCAATTAGAACGCAACACCGCAAAAGAAAATTCAAACGTT	780
Db	13230	TGGGTAAAAAATGGCGTGGATCTATTAGAAACGCAACCCGCAAAACGAAAATTCAAACGTT	13288
Oy	781	ACTGTCCGTCCTCGGGACAAATAGACCCCTACGCTAAATTGGGGTAAACCTGGCTGCCGTGCTG	840
Db	13290	ACTGTCCGTCCTCGGGACAAATAGACCCCTACGCTAAATTGGGGTAAACCTGGCTGCCGTGCTG	13349
Oy	841	AGCGCAACGATCGGCGCTCGTCAATTGTAAATTCATCGTCACAGAAATCTGTGCACCCCG	900
Db	13350	AGCGCAACGATCGGCGCTCGTCAATTGTAAATTCATCGTCACAGAAATCTGTGCACCCCG	13409
Oy	901	CACGGAATAATTAGACACGGTCTGGCAAGACAGAGAAACGTTCCCAAACTGAAAGGGA	960
Db	13410	CACGGAATAATTAGACACGGTCTGGCAAGACAGAGAAACGTTCCCAAACTGAAAGGGA	13469
Oy	961	TCGCGAAAATTTGGACCCATGATGTGTGGCTGCGGAATTAACAAAGGGCGCTGACCAAGATAGT	1020
Db	13470	TCGCGAAAATTTGGACCCATGATGTGTGGCTGCGGAATTAACAAAGGGCGCTGACCAAGATAGT	13529
Oy	1021	GAACCTGTGGAACATGTTGCGATTTGTTAAACCCGCTTGCGCTAAAGCTGCGCCGACCTCATATA	1080
Db	13530	GAACCTGTGGAACATGTTGCGATTTGTTAAACCCGCTTGCGCTAAAGCTGCGCCGACCTCATATA	13589
Oy	1081	AAAAATGGA 1089	
Db	13590	AAAAATGGA 13598	

QY	1081	AAAATGTGA	1089
Db	13590	AAAATGTGA	13598

RESULT 4

ID AAT33513 standard; DNA; 534 BP.

AC AAT33513;

DT 19-NOV-1996 (first entry)

DE Infectious laryngotracheitis virus ORF8 reverse complement.

KW Infectious laryngotracheitis virus; ILTV; herpesvirus;

ds.

OS Infectious laryngotracheitis virus.

PN W09508622-A1

PD 30-MAR-1995.

PF 16-SEP-1994; 94WO-US10628.

PR 24-SEP-1993; 9305-0126597.

PA (SYTR) SYNTRO CORP.
XX

PI Cochran MD, Wild MA;

DR WPI; 1995-139591/18

XX
XX

PT Recombinant attenuated infectious laryngotracheitis virus - for use
PT in vaccines to protect poultry from infection from the virus, also
PT methods of distinguishing between vaccinated and naturally infected
PT birds

PS Example 1; Page 115-116; 177pp; English.

CC ORF8 reverse complement (AAT33513) consists of bases 11,150 to
CC 10,617 on the reverse complement strand of the unique short
CC region (see also AAT33504) of infectious laryngotracheitis virus
CC (ILT) genomic DNA. It codes for a protein (AAM00642) of
CC approx. 19,470 mol.wt. of unknown function.

SO Sequence 4411529 BP; 758565 A; 1449983 C; 1444602 G; 758379 T; 0 other;
Query Match 3.5%; Score 38.2; DB 22; Length 4411529;
Best Local Similarity 58.3%; Pred. No. 8.9;
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
OY 29 TCCCTGCGGAGAGCTTCGACCCCTTCGCGCGATGGGAATCGTATCAGTGAATCAGC 88
DB 618034 TCCGACGCGCGACGCTGCGACCCCTCGGTGCGCGAGGCAACGATCACCCTCGGTACGCG 617975
OY 89 TCTCGGCGAGATTGACGAGATCAATCGTATCGTGGCGCTCGCGCGGAGC 143
DB 617974 TCGGCCCCAGGACGAGTGTCCAGCCGCGAGCAACCGCGCCGAGC 617920

RESULT 7
AAF26455
ID AAF26455 standard; cDNA; 2526 BP.
XX AAF26455;
XX 26-MAR-2001 (first entry)
XX D. melanogaster corkscrew DNA.
DE SHP-2; SHP-1; Src Homology-2; protein tyrosine phosphatase; mutant;
KM neoplastic disorder; obesity; angiogenesis; cancer; immune;
KM hemotopoietic; allergy; ss.
XX Drosophila melanogaster.
OS US6156551-A.
PN 05-DEC-2000.
XX 05-JUN-1998; 98US-0092443.
XX 05-JUN-1998; 98US-0092443.
PR 05-JUN-1998; 98US-0092443.
XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA (JOSL-) JOSLIN DIABETES CENT.
XX Neel BG, Shoelison S, Pluskey S, O'Reilly AM;
PI WPI; 2001-060166/07.
DR Mutant SH2 domain-containing protein tyrosine phosphatase, useful in in
XX vitro assays to screen for binding partners, inhibitors of tyrosine
PT phosphatase and for treating tyrosine phosphatase-mediated diseases
XX Disclosure; Fig 7; 161pp; English.
XX The present invention relates to an activated SH2 (Src Homology-2)
CC -domain containing protein tyrosine phosphatase (SHP-1 or SHP-2)
CC mutant with a mutation in the SH2 domain. Activated mutants of
CC SH2 domain-containing protein tyrosine phosphatases are useful in
CC in vitro assays to screen for binding partners and inhibitors of
CC the phosphatase and in the treatment of PPP-mediated diseases or
CC conditions in a mammal, including neoplastic disorders, obesity and
CC to inhibit angiogenesis. Inhibitors identified using the activated
CC mutants are useful for the treatment of cancer, immunosuppression,
CC immunostimulation, hemotopoietic stimulation and anti-allergy
CC treatment.
XX Sequence 2526 BP; 616 A; 742 C; 732 G; 436 T; 0 other;
SO

Query Match 3.5%; Score 38; DB 22; Length 2526;
Best Local Similarity 51.8%; Pred. No. 0.26;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
OY 569 TTGAGGCTATCTCAGAGCAGGAGAAAGTTGGCGCAACTGCGTACGTTGCCACGG 628
DB 1325 TTAAAGACTATCATCGACCCAGGCGTGTCTCTACCCAGCAAGTAACACGCGTGACGG 1384

OY 629 AGGCCAGACGACGAGCCGCGAGCGACAAACCCGAGCCGTCATGCAACCGGCT 688
DB 1385 ACTTTCGAGACATGCTGTGCGAGAGAAACGGGGTGATGTCATGACCAACGAGAGT 1444
OY 689 CCGAAGCTTGAAGCGGAACACTTACCTTCCCTGGCTAGAAATGC 734
DB 1445 ACGAGCGCGCGCAAAAGAAAGTGGCCCGCTACTGCGCGAGCAGCG 1490

RESULT 8
ABL06927
ID ABL06927 standard; cDNA; 2677 BP.
XX ABL06927;
XX 26-MAR-2002 (first entry)
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 15263.
XX Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ss.
XX Drosophila melanogaster.
OS WO200171042-A2.
PN 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR P-PSDB; ABB62824.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX Claim 1; SEQ ID NO 15263; 21pp + Sequence Listing; English.
PS The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABU16176-ABU130511), expressed DNA
CC sequences (ABU1840-ABU16175) and the encoded proteins
CC (ABBS737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 2677 BP; 670 A; 760 C; 749 G; 498 T; 0 other;
SO

Query Match 3.5%; Score 38; DB 23; Length 2677;
Best Local Similarity 51.8%; Pred. No. 0.27;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
OY 569 TTGAGGCTATCTCAGAGCAGGAGAAAGTTGGCGCAACTGCGTACGTTGCCACGG 628
DB 1464 TTAAAGACTATCATCGACCCAGGCGTGTCTCTACCCAGCAAGTAACACGCGTGACGG 1523
OY 629 AGGCCAGACGACGAGCCGCGAGCGCAACCCGAGCGCGCTACATGCAACCGGCGCT 688
DB 1524 ACTTTCGAGACATGCTGTGCGAGAGAAACCGCGGTGATGTCATGACCAACGAGAGT 1583

```
OY 689 CCGAAGCTGAGCGGACGACCTTACCTTCCCTGGCTGAGAAAATGG 734
ID ABL18521
DB 1584 ACGAGCGCGGCAAGAAAGTGGCCCGCTACTGCGCGGACGAGGG 1629

RESULT 9
ABL18521 standard; DNA; 2838 BP.
XX
AC ABL18521;
XX
DE 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 7036.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN MO200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001MO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Claim 1; SEQ ID NO 7036; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL10840-ABL16175) and the encoded proteins
XX (ABBS7737-ABBS72072).
XX
CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2838 BP; 711 A; 853 C; 799 G; 475 T; 0 other:

Query Match 3.5%; Score 38; DB 23; Length 2838;
Best Local Similarity 51.8%; Pred. No. 0.28;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

OY 569 TTGAGGCTATCTCAGAGCCGAGGAAAGTTGGCGCAACTGCAAGCGTAGCTTGCCACGG 628
ID ABL18520
DB 1625 TTAAAGACCTACATCCCGCCAGCGGCTGTCTGCTACCCAGCAAGTGAACGCGTGACGG 1684
OY 629 AGGCGACGAGCAGCGCGCGCGGCGAGCAACCCCGAGCGCGCTGCACTGCAACGAGCGCT 688
DB 1685 ACTTCTGAAACATGTGTGCGAGGAAACACGCGGCTGTCTGATGACCCCAAGAGT 1744
OY 689 CCGAAGCTGAGCGGACGACCTTACCTTCCCTGGCTGAGAAAATGG 734
DB 1745 ACGAGCGCGGCAAGAAAGTGGCCCGCTACTGCGCGGACGAGGG 1790

RESULT 10

ABL21135
ID ABL21135 standard; DNA; 4528 BP.
XX
AC ABL21135;
XX
DE 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 14878.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN MO200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001MO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Claim 1; SEQ ID NO 14878; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABBS7737-ABBS72072).
XX
CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4528 BP; 1311 A; 1118 C; 1051 G; 1048 T; 0 other:

Query Match 3.5%; Score 38; DB 23; Length 4528;
Best Local Similarity 51.8%; Pred. No. 0.35;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

OY 569 TTGAGGCTATCTCAGAGCCGAGGAAAGTTGGCGCAACTGCAAGCGTAGCTTGCCACGG 628
DB 1464 TTAAAGACCTACATCCCGCCAGCGGCTGTCTGCTACCCAGCAAGTGAACGCGTGACGG 1523
OY 629 AGGCGACGAGCAGCGCGCGCGGCGAGCAACCCCGAGCGCGCTGCACTGCAACGAGCGCT 688
DB 1524 ACTTCTGAAACATGTGTGCGAGGAAACACGCGGCTGTCTGATGACCCCAAGAGT 1583
OY 689 CCGAAGCTGAGCGGACGACCTTACCTTCCCTGGCTGAGAAAATGG 734
DB 1584 ACGAGCGCGGCAAGAAAGTGGCCCGCTACTGCGCGGACGAGGG 1629

RESULT 11
ABL18520
ID ABL18520 standard; DNA; 10587 BP.
XX
AC ABL18520;
XX
DE 26-MAR-2002 (first entry)
```


OS Homo sapiens.

XX WO200218632-A2.

XX 07-MAR-2002.

XX 01-SEP-2001; 2001WO-EP10074.

XX 01-SEP-2000; 2000DE-1043826.

XX 05-SEP-2000; 2000DE-1044543.

XX (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K, Guetlig D;

XX WPI; 2002-371829/40.

PT Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (1) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (11) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.

XX Sequence 813 BP; 250 A; 350 C; 99 G; 114 T; 0 other;

XX Query Match 3.5%; Score 37.6; DB 24; Length 813;

XX Best Local Similarity 49.0%; Pred. No. 0.2; Mismatches 104; Indels 0; Gaps 0;

XX Matches 100; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

XX 611 AAGGCTACGTTGCCAGGAGCCAGCAGCAACGCCGAGCGACAACCCCGAGCGCCG 670

XX 42 AAACGCCCGCGAGCCGATACAAACAGCAGCGCGCTCCCGGCTCCCGCGAGCGCA 101

XX 671 TCACGTGACCAAGCGCTCCGACCTTGAACGGAACCTTACCTTCCCTGGCTAGAAA 730

XX 102 ACCCGCGATCCCGGACCCGAAACGACACTCTCGAATCCGACTCCCGCGCGCAAA 161

XX 731 ATGGCGTGATCATTTAGAACCGACCGCAAAATTCAAACGTTACTGTGCTC 790

XX 162 ATAACTAACCCGACTACGAAACCGCGCCGAAACGAAACGAAACGAAACGCTTGC 221

XX 791 TCGGAGCAATGAGCCCTACGCTAA 814

XX 222 GCCCGCAAAAGCGCTCGACAA 245

XX Search completed: July 1, 2003, 14:40:22

XX Job time : 343 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2003, 14:16:19 ; Search time 1895 Seconds

(without alignments)
9307.063 Million cell updates/sec

Title: US-08-468-190C-1_COPY_9874_10962

Sequence: 1 ATGGCATCGCTACTTGGAAC.....CCGACTCATATATAATGCA 1089

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estin:*
3: em_estin:*
4: em_estin:*
5: em_estin:*
6: em_estin:*
7: em_estin:*
8: em_estin:*
9: em_estin:*
10: em_estin:*
11: em_estin:*
12: em_estin:*
13: em_estin:*
14: em_estin:*
15: em_estin:*
16: em_estin:*
17: em_estin:*
18: em_estin:*
19: em_estin:*
20: em_estin:*
21: em_estin:*
22: em_estin:*
23: em_estin:*
24: em_estin:*
25: em_estin:*
26: em_estin:*
27: em_estin:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	43.2	4.0	493	12	BE777414
C 2	43.2	3.9	681	9	AU076048
C 3	41.4	3.8	591	9	AU068783
C 4	41.4	3.8	765	17	AZ132997
C 5	41.4	3.8	1053	17	CNS060GTI
C 6	38.2	3.5	575	13	B1352860

C 7	37.8	3.5	925	17	CNS0091P
C 8	37	3.4	550	13	BM560053
C 9	37	3.4	675	13	BM603983
C 10	37	3.4	731	13	BM585009
C 11	36.6	3.4	409	9	AT455021
C 12	36.6	3.4	487	9	AA803404
C 13	36.6	3.4	487	9	AA841130
C 14	36.6	3.4	495	13	AT295118
C 15	36.6	3.4	538	13	B1356698
C 16	36.6	3.4	558	13	B1242364
C 17	36.6	3.4	559	9	AA941412
C 18	36.6	3.4	573	13	B1166941
C 19	36.6	3.4	588	9	AA540589
C 20	36.6	3.4	611	13	B1484946
C 21	36.6	3.4	616	13	B1353787
C 22	36.6	3.4	621	13	B1169795
C 23	36.6	3.4	625	9	AA1516547
C 24	36.6	3.4	632	9	AA948960
C 25	36.6	3.4	632	13	B1580110
C 26	36.6	3.4	648	9	AA1512258
C 27	36.6	3.4	652	9	AA820152
C 28	36.6	3.4	738	9	AA391792
C 29	36.6	3.4	742	9	AT259265
C 30	36.6	3.4	778	9	AA539582
C 31	36.6	3.3	1580	13	B1219863
C 32	36.2	3.3	1101	17	CNS0000D1
C 33	36	3.3	522	13	BM317560
C 34	36	3.3	538	12	BP292594
C 35	36	3.3	617	9	AT657056
C 36	36	3.3	624	14	BO294678
C 37	36	3.3	980	17	CNS01250
C 38	35.8	3.3	384	17	B37530
C 39	35.6	3.3	379	17	A2131897
C 40	35.6	3.3	421	9	AA693550
C 41	35.6	3.3	421	9	AA693550
C 42	35.6	3.3	421	9	AA693550
C 43	35.4	3.3	709	14	BO840906
C 44	35.4	3.3	367	13	BM448221
C 45	35.4	3.3	482	12	BF040664
			489	13	BM447647

ALIGNMENTS

RESULT 1
LOCUS BE777414/C 493 bp mRNA linear EST 20-SEP-2000
DEFINITION MY-35-A-04 PinfeastansMY Phytophthora infestans cDNA, mRNA sequence.
ACCESSION BE777414
VERSION BE777414.1 GI:10231053
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
Contact: Govers F
Laboratory of Phytopathology
Wageningen University
Bilthoven 9, P.O.Box 8025, 6700 EE, Wageningen, The Netherlands
Tel: 31 317 483 138
Fax: 31 317 483 412
Email: Francine.Govers@medew.fyto.wau.nl.
1. 493
/organism="Phytophthora infestans"
/strain="DDR7602, AI mating type"
/db_xref="taxon:4787"

BASE COUNT
ORIGIN

87 a 128 c 183 g 95 t

/clone_11b-"PinfestansMY"
/dev:string="4-week old vegetative, non-sporulating
mycelium in synthetic medium"
/lab_host="E. coli, strain DH5-alpha"
/note="vector: pSPORI1; Site_1: SalI; Site_2: NotI; Total
RNA was isolated from mycelium of P. infestans DDR602
cultured for 4 weeks in synthetic medium. EST clones were
named by their position in the microtiter plate, preceded
by the prefix MY (for mycelial) and the successive number
of the microtiter plate (e.g. MY-06-A-04)."

Query Match	4.08:	Score 43.2:	DB 12.	Length 493;
Best Local Similarity	56.28:	Pred. No. 0.22:		
Matches	81;	Conservative	0;	Mismatches 63;
				Indels 0;
				Gaps 0

548 TCGGATTTACCTTCCCTCCAACTTGGAGCTTACTCAGGACCCAGAAAGATTTGGCGCAACT 607
 159 TCCTTCTCGCAAGCGTGCATGAAGAGCCCACTCATCAAAACCTGGAGTTCCTCTG 100
 608 GCGAAGCTAGCTTTCCTCAACGAGGCCACGACAGCAACCAAGCCGACAGGCGACATCCCGCAGC 667

[illegible]

RESULT 2	LOCUS	DEFINITION	LOCUS	DEFINITION
AU076048/c	AU076048	681 bp	linear	EST 03-APR-2002
	AU076048	Rice panicle at flowering stage	Oryza sativa (Japonica	
		cultivar-group) cDNA clone E2735_6Z, mRNA sequence.		

SOURCE ORGANISM	ORGANISM
Oryza sativa (Japonica cultivar-group).	Oryza sativa (Japonica cultivar-group)
Oryza sativa (Japonica cultivar-group)	Oryza sativa (Japonica cultivar-group)
Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, Ehrhartoideae, Oryzaeae, Oryza.	

REFERENCE	1 (bases 1 to 681)
AUTHORS	Sasaki, T. and Yamamoto, K.
TITLE	Rice cDNA from panicle at flowering stage
JOURNAL	Unpublished (1996)
COMMENT	Contact: Takuji Sasaki

National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel.: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasakl@agr.affrc.go.jp, [URL: http://rgp.dna.affrc.go.jp/](http://rgp.dna.affrc.go.jp/)
PROJECT "RGP"

FEATURES	Location/Qualifiers
source	1. .681

BASE COUNT ORIGIN	141 a	171 c	191 g	174 t	4 others
/organism="Oryza sativa (japonica cultivar-group)"					
/cultivar="Nipponbare"					
/db_xref="taxon:39947"					
/clone_1="E2735_62"					
/clone_1lb="Rice panicle at flowering stage"					
/dev_stage="flowering stage"					
/note="Organ: panicle; Rice cDNA from panicle at flowering stage"					

Query Match	3.9%	Score 42;	DB 9;	Length 681;
Best Local Similarity	53.4%;	Pred. No. 0.59;		
Matches 87;	Conservative 0;	Mismatches 76;	Indels 0;	Gaps 0;

560 CCGTCCAACTTGAGGCCATATCTCAGAGACCGAGAAATGTGGCGGCACTGGCAAGCGTAC 619

Db	354	CCATCCGCGGTTTGAAGACTGCTTTTGGAAAGATGTGTGACGTGGCGCAGAGCGCTGGCCCATGT	295
Qy	620	TTGGCCAGGAGGCCAAGAGACGACGAGCGCCGAGGCGCAACCCCGACGCCCGTCATCTGCAA	679
Db	294	CGGCCACGAGGCCACCGCGCGGAGAGGAGACGACGAGCGCGCGCGGACGTTGAGCGCCT	235
Qy	680	CCAGGCGCCTCGGAACTTGAAGCGGAACACTTTACCTTCCTTCG	722
Db	234	GCAGGCGCCACCACTTCGCGCTGCTCTCTTGAAACCTTGCCTCTG	192

RESULT 3	LOCUS	DEFINITION
A0068783/c	A0068783	591 bp mRNA linear EST 01-APR-2002
	A0068783	Rice callus <i>Oryza sativa</i> (japonica cultivar group) cDNA clone C50531_102, mRNA sequence.

ACCESSION	AU068783	GI:5003634
VERSION	AU068783.1	
KEYWORDS	EST	
SOURCE	<i>Oryza sativa</i> (japonica cultivar-group)	
ORGANISM	<i>Oryza sativa</i> (japonica cultivar-group)	

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 591)
Yamamoto, K. and Sasaki, T.
Rice cDNA from callus (1998)

Euariyot; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharitoidae; Oryzaceae; Oryza.
Rice cDNA from callus (1998)

JOURNAL COMMENT
Unpublished (1998)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel.: 81-298-38-7441
Fax: 81-298-38-7458
Email: tsasaki@agr.affrc.go.jp, URL: <http://rjp.dna.affrc.go.jp/>
PROJECT: 'RGP'

FEATURES
SOURCE

BASE COUNT	ORIGIN
137 a	137 a
139 c	139 c
159 g	159 g
155 t	155 t
1 others	1 others

Query Match	3.8%	Score	41.4	DB	9	Length	591
Best Local Similarity	53.4%	Pred. No.	0.81				
Matches	87	Conservative	0	Mismatches	76	Indels	0
						Gaps	0

QY	560	CCGCCCAACTTGAAGGCTATCTCAGGACCGAGAAAGTTGGCGCAACTGCGCAAGCTACG	61.9
QY	260	CCATCTCGGGTTTCAGAGCTGCGTTTGGAAAGATGTGACGTGGGCGCGCTGCGCCCATGT	201
QY	620	TTGGCAGAGAGGCGCAGACGACGACGCGCGAGGCGCAACCCCGACGCGCTCTACTGCA	67.9
DB	200	CGGCACCGAGCCCGACCGCGCGGAGGAGCGAACAAGAGCGCGCGACGTTGAGCGCT	14.1
QY	680	CCAGCGCTCGAACTTGAAGCGGAACACTTTATCTTCCCTG	72.2
DB	140	GCAAGCGCACCACTTCGCGGTGCTCTCTGCACTTCGCGCTG	98

RESULT 4	
A213297	
LOCUS	A213297
DEFINITION	765 bp DNA linear GSS 02-JUN-2000
	OSJNB0108G22f CUGI Rice BAC Library (EcorI) <i>Oryza sativa</i> genomic
	clone OSJNB0108G22f, DNA sequence.

ACCESSION	AZ132997
VERSION	AZ132997.1 GI:8211738
KEYWORDS	GSS.
SOURCE	Oryza sativa.
ORGANISM	Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
REFERENCE	Euhartoidae; Oryzaceae; Oryza.
AUTHORS	Wing R.A. and Dean R.A.
TITLE	A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL	Unpublished (1998)
COMMENT	Contact: Wing RA Clemson University Genomics Institute 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: rwing@clemson.edu Seq primer: GTTAAAGCACGCGCCATGTG Class: BAC ends High quality sequence start: 31 High quality sequence stop: 580. Location/Qualifiers
FEATURES	
source	1..765 /organism="Oryza sativa" /strain="Japanica" /cultivar="Nipponbare" /db_xref="taxon:4530" /clone="OSUBINO108622F" /clone_id="CGR Rice BAC library (ECORI)" /library_type="Leaf Rice DH10B" /lab_host="E. coli DH10B" /note="Vector: pBlacindigo; Site_1: EcoRI; Site_2: EcoRI; Rice is the most important food crop in the world. Half of the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 MbP (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa. Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EEST Resource Center (www.genome.clemson.edu)."
BASE COUNT	186 a 228 c 204 g 145 t 2 others
ORIGIN	
Query Match	3.8%; Score 41.4; DB 17; Length 765;
Best Local Similarity	53.4%; Pred. No. 0.94; 76; Indels 0; Gaps 0;
Matches	87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
D1	560 CCGTCACACTTGAGGGCCCTATTCAAGACCAGGAAGTTGGCGCAACTGGCAGCGTAGC 619
D2	
D3	280 CCATCGGGGTTCAAGAGTGCAGTTGGAGAAGAGTGTGAGCGGCGACGCCGCTGCACCATGT 339
D4	620 TTGCGACGAGGAGCGAGAGCAGACCGCGCGCGAGCGAACAACCCGAGCGCCCTGACACGCAA 679
D5	
D6	340 CGGCCACGACCCACCGCCCGCGAGACAGCAAGAACAGGCGCGCGACGTTGAGCGCCT 399
D7	680 CCAAGCGCTCCGCAACTTGAAAGCGGAACCTTACTCTTCCCTG 722
D8	
D9	400 GGAGGCCCACCACTTCGCGGTGCCTCTCTGCACACTTCGCGCTG 442

[illegible]

Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)"

BASE COUNT 182 a 173 c 244 g 132 t
ORIGIN

Query Match 3.4%; Score 37; DB 13; Length 731;

Best Local Similarity 53.0%; Pred. No. 17;

Matches 79; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Db 543 CACCTTCGATTACCGACGCTTACCTGATCTACGAGCAGGAAGTTGGCG 602

459 CACCTTCGCTTTGTTTATACCACTCTTGTCCGCCACTGTAAGCCGCGTTGGT 400

Qy 603 CAACTGGCAAGGATGATGTCACGAGCAGACGACGCGGACGACACACCC 662

Db 399 TACCACTCTCTGTGTAGCCGCCGCCACACACGACGCGGCTGTACAC 340

Qy 663 GACGCGGCTACTGACACGACGCGCTCCG 691

Db 339 CGTACCGCCACCGTCACACGCGCTCCG 311

RESULT 11

LOCUS A1455021/c 409 bp mRNA linear EST 19-APR-2001

DEFINITION LD04335.5prime LD Drosophila melanogaster embryo Bluescript

Drosophila melanogaster cDNA clone LD04335 5prime, mRNA sequence.

ACCESSION A1455021

VERSION A1455021.1 GI:4298123

KEYWORDS EST.

SOURCE fruit fly.

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 409)

Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,

Lewis, S. and Rubin, G.M.

BDGP/HHMI Drosophila EST Project

Unpublished (2001)

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu

Plate: 43 row: C column: 11

High quality sequence stop: 363.

Location/Qualifiers

1..409

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone="LD04335"

/clone_1lb="LD Drosophila melanogaster embryo Bluescript"

/sex="male and female"

/dev_stage="0 to 24 hours mixed stage embryonic"

/lab_host="SOLR"

/note="Organ: embryo; Vector: Bluescript SK; Site: 1; EcORI

Synthesis kit. Oligo dt-primed and directionally cloned at

EcORI and XhoI in Bluescript SK(+/-)"

BASE COUNT 127 a 88 c 128 g 66 t

ORIGIN

Query Match 3.4%; Score 36.6; DB 9; Length 409;

Best Local Similarity 62.6%; Pred. No. 16;

Matches 57; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 29 TCCTTGGCGGACGCTTCGACCCCTTCGCGGATGGATCGTACACTGGAATCAG 88

Db 281 TCGTTTTCGCTTCTTTCGCGGATGGATCGTACACTGGAATCGTCC 222

Qy 89 TCTCCGACGATGACGATCAGTACATCT 119

Db 221 TCTCGATGCGATGTCGTGATCAGATCTT 191

RESULT 12

LOCUS AA803404/c 487 bp mRNA linear EST 23-APR-2001

DEFINITION GM10781.5prime GM Drosophila melanogaster ovary Bluescript

Drosophila melanogaster cDNA clone GM10781 5prime, mRNA sequence.

ACCESSION AA803404

VERSION AA803404.1 GI:2872523

KEYWORDS EST.

SOURCE fruit fly.

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 487)

Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,

Lewis, S. and Rubin, G.M.

BDGP/HHMI Drosophila EST Project

Unpublished (2001)

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu

Plate: 107 row: G column: 9

High quality sequence stop: 380.

Location/Qualifiers

1..487

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone="GM10781"

/clone_1lb="GM Drosophila melanogaster ovary Bluescript"

/sex="female"

/dev_stage="newly eclosed females: germarium-stage 6"

/lab_host="SOLR"

/note="Organ: ovary; Vector: Bluescript SK; Site: 1; EcORI;

Site: 2; XhoI: Constructed using Stratagene ZAP-cDNA

Synthesis kit. Oligo dt-primed and directionally cloned at

EcORI and XhoI in Bluescript SK(+/-)"

BASE COUNT 143 a 104 c 151 g 88 t 1 others

ORIGIN

Query Match 3.4%; Score 36.6; DB 9; Length 487;

Best Local Similarity 62.6%; Pred. No. 18;

Matches 57; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 29 TCCTTGGCGGACGCTTCGACCCCTTCGCGGATGGATCGTACACTGGAATCAG 88

Db 312 TCGTTTTCGCTTCTTTCGCGGATGGATCGTACACTGGAATCGTCC 253

Qy 89 TCTCCGACGATGACGATCAGTACATCT 119

Db 252 TCTCGATGCGATGTCGTGATCAGATCTT 222

RESULT 13

LOCUS AA941430/c 487 bp mRNA linear EST 19-APR-2001

DEFINITION LD25565.5prime LD Drosophila melanogaster embryo POT2 Drosophila

melanogaster cDNA clone LD25565 5prime, mRNA sequence.

ACCESSION AA941430

VERSION AA941430.1 GI:3101343

KEYWORDS EST.

SOURCE fruit fly.

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 487)

Query Match	3.48;	Score 36.6;	DB 13;	Length 538;
Best Local Similarity	62.68;	Pred. NO. 19;		
Matches 57; Conservative	0;	Mismatches 34;	Indels 0;	Gaps 0;

[illegible]

```

Search completed: July 1, 2003, 16:03:52
Job time : 1903 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2003, 14:30:58 ; Search time 70 Seconds
(without alignments) 4771.014 Million cell updates/sec

Title: US-08-468-190C-1_COPY_9874_10962

Perfect score: 1089

Sequence: 1 ATGCATCGCTACTTGGAAC.....CCGACTCAATTAATAATGTGA 1089

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 segs, 15338381 residues

Minimum number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents -NA:*
1: /cgn2_6/ptodata/1/ins/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ins/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ins/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ins/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ins/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ins/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1089	100.0	13473	5 PCT-US96-03916-1	Sequence 1, Appl
2	1087.4	99.9	18912	5 PCT-US96-03916-59	Sequence 59, Appl
3	346	31.8	534	5 PCT-US96-03916-18	Sequence 18, Appl
4	38.2	3.5	4403765	4 US-09-103-840A-2	Sequence 2, Appl
5	38.2	3.5	4411529	4 US-09-103-840A-1	Sequence 1, Appl
6	35.2	3.2	19718	4 US-08-961-527-99	Sequence 99, Appl
7	34.6	3.2	50341	1 US-08-247-901C-1	Sequence 1, Appl
8	34.6	3.2	50341	2 US-09-075-904-1	Sequence 1, Appl
9	34.6	3.2	52297	4 US-09-426-436-1	Sequence 1, Appl
10	34.6	3.0	52297	4 US-08-705-557-1	Sequence 1, Appl
11	32.2	3.0	4403765	4 US-09-103-840A-2	Sequence 2, Appl
12	32.2	3.0	4411529	4 US-09-103-840A-1	Sequence 1, Appl
13	31.8	2.9	1145	3 US-08-809-103B-7	Sequence 7, Appl
14	31.8	2.9	1148	3 US-08-809-103B-1	Sequence 1, Appl
15	31.8	2.9	1150	3 US-08-809-103B-3	Sequence 3, Appl
16	31.8	2.9	7218	3 US-08-809-103B-5	Sequence 5, Appl
17	31.8	2.9	7218	3 US-08-232-653-14	Sequence 14, Appl
18	31.6	2.9	843	3 US-08-513-974B-375	Sequence 375, Appl
19	31.6	2.9	25165	4 US-09-453-702B-39	Sequence 39, Appl
20	30.8	2.8	948	4 US-09-134-001C-118	Sequence 118, Appl
21	30.8	2.8	1728	1 US-08-477-097-1	Sequence 1, Appl
22	30.8	2.8	1728	1 US-08-878-957-1	Sequence 1, Appl
23	30.8	2.8	1752	1 US-08-477-097-15	Sequence 15, Appl
24	30.8	2.8	1752	1 US-08-477-097-27	Sequence 27, Appl
25	30.8	2.8	1752	2 US-08-878-957-15	Sequence 15, Appl
26	30.8	2.8	1752	2 US-08-878-957-27	Sequence 27, Appl
27	30.4	2.8	3467	1 US-07-745-206A-12	Sequence 12, Appl

28	30.4	2.8	5467	2 US-08-311-363-12	Sequence 12, Appl
29	30	2.8	843	4 US-08-339-214-7	Sequence 7, Appl
30	30	2.8	843	4 US-09-006-755B-2	Sequence 2, Appl
31	30	2.8	846	3 US-08-815-190A-1	Sequence 1, Appl
32	30	2.8	924	4 US-08-339-214-33	Sequence 33, Appl
33	30	2.8	924	4 US-08-339-214-34	Sequence 34, Appl
34	30	2.8	924	4 US-08-339-214-35	Sequence 35, Appl
35	30	2.8	972	4 US-09-479-524-7	Sequence 7, Appl
36	30	2.8	972	4 US-09-479-524-7	Sequence 7, Appl
37	30	2.8	1790	2 US-08-810-453-1	Sequence 1, Appl
38	30	2.8	1841	5 PCT-US95-00362-1	Sequence 1, Appl
39	30	2.8	1890	4 US-09-290-640-24	Sequence 24, Appl
40	30	2.8	2432	4 US-08-339-214-29	Sequence 29, Appl
41	30	2.8	3423	4 US-08-471-112A-2	Sequence 2, Appl
42	30	2.8	5430	3 US-09-012-515A-11	Sequence 11, Appl
43	30	2.8	5430	3 US-08-360-144A-11	Sequence 11, Appl
44	30	2.8	5430	3 US-09-012-504A-11	Sequence 11, Appl
45	30	2.8	7653	4 US-08-471-112A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
PCT-US96-03916-1
Sequence 1, Application PC/TUS9603916
GENERAL INFORMATION:
APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
NUMBER OF SEQUENCES: 72
TITLE OF INVENTION: AND USES THEREOF
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 391-0525
TELEFAX: (212) 278-0400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 13473 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1059..2489
FEATURE:
NAME/KEY: CDS
LOCATION: 2575..4107
FEATURE:

NAME/KEY: CDS
LOCATION: 4113..4445
FEATURE:
NAME/KEY: CDS
LOCATION: 4609..5487
FEATURE:
NAME/KEY: CDS
LOCATION: 5697..8654
FEATURE:
NAME/KEY: CDS
LOCATION: 9874..10962
FEATURE:
NAME/KEY: CDS
LOCATION: 11159..12658
FEATURE:
NAME/KEY: CDS
LOCATION: 12665..13447
PCT-US96-03916-1

Query Match 100.0%; Score 1089; DB 5; Length 13473;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1089; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCATCGCTACTTGGAACTCTGCTCTCTCCCGGAGCTCGACCCCTTGGCGCG 60
DB 9874 ATGGCATCGCTACTTGGAACTCTGCTCTCTCCCGGAGCTCGACCCCTTGGCGCG 9933
QY 61 ATGGGAATCGATGATCTGGAATCATGCTCTCCGCAAGATTGACAGATCATGCTG 120
DB 9934 ATGGGAATCGATGATCTGGAATCATGCTCTCCGCAAGATTGACAGATCATGCTG 9993
QY 121 ATGCTCGGCTCGCCCGGAACTGATCAATTCAGTATTTTCATGCTGGCCAG 180
DB 9994 ATGCTCGGCTCGCCCGGAACTGATCAATTCAGTATTTTCATGCTGGCCAG 10053
QY 181 AGACCCGCAAAACCTCTAGTGAACCGTCCGCGCTTGGCTGATATTAACAAC 240
DB 10054 AGACCCGCAAAACCTCTAGTGAACCGTCCGCGCTTGGCTGATATTAACAAC 10113
QY 241 CAGTGTACAGAACTTAGCGAGAGCGCTTTGAAAATTGACATCATGCTCTCT 300
DB 10114 CAGTGTACAGAACTTAGCGAGAGCGCTTTGAAAATTGACATCATGCTCTCTCT 10173
QY 301 GTTTTGTGCGCTGTAAGTACGAGTACAGCTTCTCCGCTCGGAACAGACTAACCGGA 360
DB 10174 GTTTTGTGCGCTGTAAGTACGAGTACAGCTTCTCCGCTCGGAACAGACTAACCGGA 10233
QY 361 CCTCCACACCGCTTAACTGCTACATACGAATCTCTGCTCGGAACGACAGCGGATGTT 420
DB 10234 CCTCCACACCGCTTAACTGCTACATACGAATCTCTGCTCGGAACGACAGCGGATGTT 10293
QY 421 TACGTAATGTTGCTGCTAGAGACACCAAAAGAACCTTTCGCTTCGATCACTA 480
DB 10294 TACGTAATGTTGCTGCTAGAGACACCAAAAGAACCTTTCGCTTCGATCACTA 10353
QY 481 TCGGTATCAATTCGGAACACCGCGGCACTCGGAGACTTATTCGAAGCTTCGT 540
DB 10354 TCGGTATCAATTCGGAACACCGCGGCACTCGGAGACTTATTCGAAGCTTCGT 10413
QY 541 CGGACCTTGGATTAATCTACCGTTCGATTCGAGCCCTATCTCAGAGACCGAGAAAGTTGG 600
DB 10414 CGGACCTTGGATTAATCTACCGTTCGATTCGAGCCCTATCTCAGAGACCGAGAAAGTTGG 10473
QY 601 CGCAACTGCAAGGCTAGTTGCGACGAGGACGACGACGACGAGCGGAGGCAAC 660
DB 10474 CGCAACTGCAAGGCTAGTTGCGACGAGGACGACGACGACGAGCGGAGGCAAC 10533
QY 661 CGGACGCGCTCACTGCAACGCGCTCGGAACCTTGAAGCGGAACCTTACCTTCC 720
DB 10534 CGGACGCGCTCACTGCAACGCGCTCGGAACCTTGAAGCGGAACCTTACCTTCC 10593
QY 721 TGGCTAGAAAATGGCGTGATCATTAAGAACGACGACCGCAAGAAATTCACAGTT 780
DB 10594 TGGCTAGAAAATGGCGTGATCATTAAGAACGACGACCGCAAGAAATTCACAGTT 10653

DB 10594 TGGCTAGAAAATGGCGTGATCATTAAGAACGACGACCGCAAGAAATTCACAGTT 10653
QY 781 ACTGTCGCTCGGAGAACGCTTACGCTAATTTGGGCTAACGCTGCGCTG 840
DB 10654 ACTGTCGCTCGGAGAACGCTTACGCTAATTTGGGCTAACGCTGCGCTG 10713
QY 841 AGCGCAAGATCGGCTCGTCTGATTTGATTTTCATCTGCAACGAAACATGTGACCCCG 900
DB 10714 AGCGCAAGATCGGCTCGTCTGATTTGATTTTCATCTGCAACGAAACATGTGACCCCG 10773
QY 901 CACCGAAATTTAGACAGGCTTCCGCAAGAGACGACGAAAGCTTCCCAACTAGAAAGGAA 960
DB 10774 CACCGAAATTTAGACAGGCTTCCGCAAGAGACGACGAAAGCTTCCCAACTAGAAAGGAA 10833
QY 961 TCGGAAATTTGACCATGCTTGGCGGAAATTAACAAGGCGGTGACCAAGATAGT 1020
DB 10834 TCGGAAATTTGACCATGCTTGGCGGAAATTAACAAGGCGGTGACCAAGATAGT 10893
QY 1021 GAACCTGGAACGCTTGGGATTTGTAACCCGCTGCGCTTACGCTGCGCCGACTCAATA 1080
DB 10894 GAACCTGGAACGCTTGGGATTTGTAACCCGCTGCGCTTACGCTGCGCCGACTCAATA 10953
QY 1081 AAAATGGA 1089
DB 10954 AAAATGGA 10962

RESULT 2
PCT-US96-03916-59
Sequence 59, Application PC/TUS9603916
GENERAL INFORMATION:
APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03916
FILING DATE: 23-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 18912 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N
ANTI-SENSE: N
FEATURE:
NAME/KEY: CDS

OY 611 AACG 615
DB 18485 AACG 18481

RESULT 8

US-09-075-904-1/c
Sequence 1, Application US/09075904
Patent No. 5994137
GENERAL INFORMATION:
APPLICANT: Jacobs, et al.
TITLE OF INVENTION: L5 SHUTTLE PHASMIDS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Epstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/075.904
FILING DATE: May 11, 1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/247.901
FILING DATE: May 23, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Elizabeth A
REGISTRATION NUMBER: 39.911
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 50341
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: L5 shuttle phasmid sequence
HYPOTHETICAL: NO
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: L5 mycobacteriophage
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL LINE:
ORGANELLE:
IMMEDIATE SOURCE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION: No. 5994137e
AUTHORS:
TITLE:
JOURNAL:

VOLUME:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-09-075-904-1

Query Match 3.2% Score 34.6; DB 2; Length 50341;
Best Local Similarity 49.2%; Pred. No. 2.9;
Matches 91; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

OY 431 TTCCGCTAGACGACCAAGACCAATTCAGCTTCGCGATCCAACTATCGGTATC 490
DB 18665 TTCCGCTGTGATGAGATGACACCCCAAGACATGTCTCGCTCAGGGCTCGGTGGAC 18606
OY 491 AATTGCGACACCGCGCGGAGCTGCGGACTCTATTCACAGGCTTCGTGCACTTCG 550
DB 18605 AAGTAGCGAGCGCGCGGAGTACTCCCGGCTTCAACACGACCTTGACGGAGGGTCG 18546
OY 551 GATTACCTACCGCTCCAACTTATGAGGCTTATCTCAGACCGAGGAAGTTGGGCAACTGGC 610
DB 18545 TAGAACAACCCCTCCAGCTGTGCGCAGTACACACCCCTGTCACCGGTCGCAAGTTG 18486
OY 611 AACG 615
DB 18485 AACG 18481

RESULT 9

US-09-426-436-1/c
Sequence 1, Application US/09426436
Patent No. 6225066
GENERAL INFORMATION:
APPLICANT: William R. Jacobs, Jr.
APPLICANT: Barry R. Bloom
TITLE OF INVENTION: MYCOBACTERIAL SPECIES-SPECIFIC
TITLE OF INVENTION: REPORTER MYCOBACTERIOPHAGES
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Epstein
STREET: 90 Park Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (ASCII)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/426.436
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/705.557
FILING DATE:
APPLICATION NUMBER: US/08/057.531
FILING DATE:
APPLICATION NUMBER: 07/833.431
FILING DATE: February 7, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Pasqualini, Patricia A.
REGISTRATION NUMBER: 34,894
REFERENCE/DOCKET NUMBER: 96700/238
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 1:

[illegible]

RESULT 11

```

US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
;
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
US-09-103-840A-2

```

RESULT 12

US-09-103-840A-1
 : Sequence 1, Application US/09103840A
 : Patent No. 6294328
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: FLEISCHMAN, Robert D.
 : APPLICANT: WHITE, Owen R.
 : APPLICANT: FRASER, Claire M.
 : APPLICANT: VENTER, John C.
 : TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 : TITLE OF INVENTION: TUBERCULOSIS
 : FILE REFERENCE: 24366-20007.00
 : CURRENT APPLICATION NUMBER: US/09/103,840A
 : CURRENT FILING DATE: 1998-06-24
 : NUMBER OF SEQ ID NOS: 2

```

: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 4411529
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: OTHER INFORMATION: H37Rv
: US-09-103-840A-1

```

RESULT 13

```

US-08-809-103B-7
Sequence 7, Application US/08809103B
Patent No. 6133505
GENERAL INFORMATION:
APPLICANT: GRONENBORN, Bruno
TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: YOUNG & THOMPSON
STREET: 745 South 23rd Street
City: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,103B
FILING DATE: 17-MAR-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 94.11040
FILING DATE: 15-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR95/01192
FILING DATE: 15-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 521-2297
TELEFAX: (703) 685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1145 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS

```

LOCATION: 1..1077
US-08-809-103B-7

Query Match
Best Local Similarity 50.3%; Pred. No. 4;
Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 856 CTCGCTATTGTAATTCATTCGTACACGAAGAAATGTGCACCCCGACCGAAATTTAGAC 915
DB 172 CTCATATTTCTATTCATTCATTCGAAGAAATTTAGCAATTTGACCAATCAACGATTTCTGCAC 231

QY 916 ACGGTCTGCAAGACGAGCAAGAACTGCCAACTAGAAAGGAATCGCAAAATTTGA 975
DB 232 CTGGTATCCCCAACGAGTCAGCAATTCATCCGAACTTCAGGAGCTAATCGAGC 291

QY 976 CCCATGTTGCTGCGCAAAATTAACAAAGGCGCTGA 1010
DB 292 TCCGACGTCAAGTCTATATCGAAGCAAGCGAGA 326

RESULT 14
US-08-809-103B-1
; Sequence 1, Application US/08809103B
; Patent No. 6133505
; GENERAL INFORMATION:
; APPLICANT: GRONENBORN, Bruno
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 17-MAR-1997
; APPLICATION NUMBER: US/08/809,103B
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; FILING DATE: 15-SEP-1994
; APPLICATION NUMBER: FR 94.11040
; PRIOR APPLICATION DATA:
; FILING DATE: 15-SEP-1994
; APPLICATION NUMBER: WO PCT/FR95/01192
; FILING DATE: 15-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: US94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEFAX: (703) 685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1148 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1077
US-08-809-103B-1

Query Match
Best Local Similarity 50.3%; Pred. No. 4;
Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 856 CTCGCTATTGTAATTCATTCGTACACGAAGAAATGTGCACCCCGACCGAAATTTAGAC 915
DB 172 CTCATATTTCTATTCATTCATTCGAAGAAATTTAGCAATTTGACCAATCAACGATTTCTGCAC 231

QY 916 ACGGTCTGCAAGACGAGCAAGAACTGCCAACTAGAAAGGAATCGCAAAATTTGA 975
DB 232 CTGGTATCCCCAACGAGTCAGCAATTCATCCGAACTTCAGGAGCTAATCGAGC 291

QY 976 CCCATGTTGCTGCGCAAAATTAACAAAGGCGCTGA 1010
DB 292 TCCGACGTCAAGTCTATATCGAAGCAAGCGAGA 326

RESULT 15
US-08-809-103B-3
; Sequence 3, Application US/08809103B
; Patent No. 6133505
; GENERAL INFORMATION:
; APPLICANT: GRONENBORN, Bruno
; TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 17-MAR-1997
; APPLICATION NUMBER: US/08/809,103B
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; FILING DATE: 15-SEP-1994
; APPLICATION NUMBER: FR 94.11040
; PRIOR APPLICATION DATA:
; FILING DATE: 15-SEP-1995
; APPLICATION NUMBER: WO PCT/FR95/01192
; FILING DATE: 15-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: US94AL CNR TOM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-2297
; TELEFAX: (703) 685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1150 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1077
US-08-809-103B-3

Query Match
Best Local Similarity 50.3%; Pred. No. 4;
Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 856 CTCGCTATTGTAATTCATTCGTACACGAAGAAATGTGCACCCCGACCGAAATTTAGAC 915
DB 172 CTCATATTTCTATTCATTCATTCGAAGAAATTTAGCAATTTGACCAATCAACGATTTCTGCAC 231

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2003, 15:32:09 ; Search time 226 Seconds
(Without alignments)
7152.939 Million cell updates/sec

Title: US-08-468-190c-1_COPY_9874_10962

Sequence: 1 ATGCATCGCTACTGAC.....CCAGCTCAATAAATGGA 1089

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1055720 seqs, 74222136 residues

otal number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

Published Applications NA: *
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/1/pubpna/PC1_NEM_PUB.seq: *
3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq: *
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq: *
6: /cgn2_6/ptodata/1/pubpna/PC1US_PUBCOMB.seq: *
7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq: *
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq: *
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq: *
11: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq: *
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq: *
13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq: *
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1089	100.0	13473	9	US-09-994-064-1
2	1087.4	99.9	3605	10	US-09-881-457A-3
3	1087.4	99.9	18913	9	US-09-994-064-59
4	346	31.8	534	9	US-09-994-064-18
5	346	31.3	1049	9	US-10-123-155-358
6	35	3.2	653	9	US-10-184-634-402
7	35	3.2	653	9	US-10-184-634-402
8	34.6	3.2	475	9	US-10-040-739-899
9	34.6	3.1	493	9	US-09-918-995-21276
10	34.2	3.1	802	9	US-10-184-634-312
11	34.2	3.1	802	9	US-10-184-634-312
12	34	3.1	2855	9	US-10-125-540-594
13	34	3.1	2855	10	US-09-764-870-594
14	33.2	3.0	777	9	US-10-022-832-1
15	33.4	3.0	491	10	US-09-664-761-6494
16	32.4	3.0	869	9	US-10-157-031-164
17	32.4	3.0	3024	10	US-09-833-381-889
18	31.8	2.9	1223	10	US-09-815-242-9609
19	31.6	2.9	25105	9	US-10-114-170-39

C 20	31.4	2.9	941	9	US-10-123-155-464	Sequence 464, App
C 21	31.2	2.9	458	9	US-10-184-634-546	Sequence 546, App
C 22	31.2	2.9	458	9	US-10-184-634-546	Sequence 546, App
C 23	31.2	2.9	4166	10	US-09-782-980-73	Sequence 73, App1
C 24	31.2	2.9	88421	9	US-09-976-059-1	Sequence 1, App1
C 25	31	2.8	730	9	US-10-226-489-5	Sequence 5, App1
C 26	31	2.8	1107	10	US-09-974-300-1003	Sequence 1003, App
C 27	31	2.8	1944	9	US-09-943-123-18	Sequence 18, App1
C 28	31	2.8	1944	10	US-10-123-155-296	Sequence 1735, App
C 29	30.8	2.8	1086	10	US-09-815-242-6138	Sequence 296, App
C 30	30.8	2.8	1111	9	US-09-983-446A-1	Sequence 6138, App
C 31	30.8	2.8	2322	9	US-09-983-446A-1	Sequence 2, App1
C 32	30.8	2.8	335	9	US-10-184-634-174	Sequence 174, App
C 33	30.6	2.8	335	9	US-10-184-634-174	Sequence 174, App
C 34	30.6	2.8	405	9	US-09-918-995-3924	Sequence 3924, App
C 35	30.6	2.8	405	9	US-10-184-634-442	Sequence 442, App
C 36	30.6	2.8	440	9	US-10-184-634-442	Sequence 442, App
C 37	30.6	2.8	671	9	US-10-184-634-346	Sequence 346, App
C 38	30.6	2.8	671	9	US-10-184-634-346	Sequence 346, App
C 39	30.6	2.8	1024	9	US-10-123-155-198	Sequence 198, App
C 40	30.6	2.8	1323	10	US-09-976-740-45	Sequence 4076, App
C 41	30.6	2.8	1614	9	US-10-023-529-45	Sequence 45, App1
C 42	30.6	2.8	1614	12	US-10-023-529-45	Sequence 45, App1
C 43	30.6	2.8	1614	12	US-10-023-529-45	Sequence 45, App1
C 44	30.6	2.8	2051	10	US-09-917-800A-1325	Sequence 1325, App
C 45	30.6	2.8	2051	10	US-09-917-800A-1325	Sequence 1325, App

ALIGNMENTS

RESULT 1
US-09-994-064-1
Sequence 1, Application US/09994064
Publication No. US20030082788A1
GENERAL INFORMATION:
APPLICANT: Wild, Martha A.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/994,064
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/468,190
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 261-0525
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 13473 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1059..2489
FEATURE:
NAME/KEY: CDS
LOCATION: 2575..4107
FEATURE:
NAME/KEY: CDS
LOCATION: 4113..4445
FEATURE:
NAME/KEY: CDS
LOCATION: 4609..5487
FEATURE:
NAME/KEY: CDS
LOCATION: 5697..8654
FEATURE:
NAME/KEY: CDS
LOCATION: 9874..10962
FEATURE:
NAME/KEY: CDS
LOCATION: 11159..12658
FEATURE:
NAME/KEY: CDS
LOCATION: 12665..13447
US-09-994-064-1

Query Match 100.0%; Score 1089; DB 9; Length 13473;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1089; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCAGCTGCTACTGGAATCTGGCTCTCCTTGGCGGAGCGCTCGACCCCTTGGCGCG 60
DB 9874 ATGGCAGCTGCTACTGGAATCTGGCTCTCCTTGGCGGAGCGCTCGACCCCTTGGCGCG 9933
QY 61 ATGGGAATCGTGATCAGTGAATATCAGTCTCCCGCAGAGATTGAGACGATCAGTCTG 120
DB 9934 ATGGGAATCGTGATCAGTGAATATCAGTCTCCCGCAGAGATTGAGACGATCAGTCTG 9993
QY 121 ATGCTGCGGCTCGCGCCGAGAGCTACAAATTCAGTGCAGTATTTTCATGCTGGCCAG 180
DB 9994 ATGCTGCGGCTCGCGCCGAGAGCTACAAATTCAGTGCAGTATTTTCATGCTGGCCAG 10053
QY 181 AGACCCCAAAACCTCTAGAGAACCGTCCGCGTTCGGTTCGATATTAACAAC 240
DB 10054 AGACCCCAAAACCTCTAGAGAACCGTCCGCGTTCGGTTCGATATTAACAAC 10113
QY 241 CAGTGTACGAGAACTTACGAGAGCGCTTTGAAAATTCGACTCATGATCGTCTCT 300
DB 10114 CAGTGTACGAGAACTTACGAGAGCGCTTTGAAAATTCGACTCATGATCGTCTCT 10173
QY 301 GTTTTGTGCGCTGTAAGTAGACGAGTACAGTTCCTCGGCTGGAACAGACTAACCGGA 360
DB 10174 GTTTTGTGCGCTGTAAGTAGACGAGTACAGTTCCTCGGCTGGAACAGACTAACCGGA 10233
QY 361 CTTCCACACCCGTTAAAGCTACATATAGAAATCTCTGTCGGAAGACAGCGGATGTT 420
DB 10234 CTTCCACACCCGTTAAAGCTACATATAGAAATCTCTGTCGGAAGACAGCGGATGTT 10293
QY 421 TACCTAATTTGTTGGCTAGAGACAGCAACCAAAACCATGACGTTTCGATCAACTA 480
DB 10294 TACCTAATTTGTTGGCTAGAGACAGCAACCAAAACCATGACGTTTCGATCAACTA 10353
QY 481 TCGGTATATCAATTCGGAACACCGCGAGCTCGGAGCTATTCACAAAGCTTCGCT 540
DB 10354 TCGGTATATCAATTCGGAACACCGCGAGCTCGGAGCTATTCACAAAGCTTCGCT 10413
QY 541 CGCAGCTTCGATTAACCTACCGTCAACTTGAAGCTATTCAGAGACGAGAAAGTTG 600
DB 10414 CGCAGCTTCGATTAACCTACCGTCAACTTGAAGCTATTCAGAGACGAGAAAGTTG 10473
QY 601 CGCAACTGGCAAGGCTGTTGCCACGAGGCGCACGACGACCAAGCGCGGAGCCAGCAAC 660

DB 10474 CGCAACTGGCAAGGCTGTTGCCACGAGGCGCACGACGACAGCGCGGAGCCGAGCCAGCAAC 10533
QY 661 CCGAGCGCCCTCAGTGAACAGAGCGCTCCCAACTGTAAGGGAACACTTTACCTTCC 720
DB 10534 CCGAGCGCCCTCAGTGAACAGAGCGCTCCCAACTGTAAGGGAACACTTTACCTTCC 10593
QY 721 TGGCTAGAAAATGGCGTGAATCATATTCAGAACCGACCGCAACGAAATTCAAAGCT 780
DB 10594 TGGCTAGAAAATGGCGTGAATCATATTCAGAACCGACCGCAACGAAATTCAAAGCT 10653
QY 781 ACTGTCGCTTCGGGACAAATAGCCCTTAAATGGGGTAAACCGTGGCTGCTG 840
DB 10654 ACTGTCGCTTCGGGACAAATAGCCCTTAAATGGGGTAAACCGTGGCTGCTG 10713
QY 841 AGCGCAAGATCGGCTGCTATGTATTTTCATTCGTCACAGAAATATGTCACCCG 900
DB 10714 AGCGCAAGATCGGCTGCTATGTATTTTCATTCGTCACAGAAATATGTCACCCG 10773
QY 901 CACGGAATATAGACAGCGTCTCGCAAGACGAGAAAGCTTCCCAACTAGAAAGGAA 960
DB 10774 CACGGAATATAGACAGCGTCTCGCAAGACGAGAAAGCTTCCCAACTAGAAAGGAA 10833
QY 961 TCGGAAAATTTGACACCATGCTGCGCAATTAACAAGGCGCTGACCCAGATAGT 1020
DB 10834 TCGGAAAATTTGACACCATGCTGCGCAATTAACAAGGCGCTGACCCAGATAGT 10893
QY 1021 GAACCTTGGGAAGTGGTGGATTTGTAACCGCTGCGTACGCTGCGCGAGCTCAATA 1080
DB 10894 GAACCTTGGGAAGTGGTGGATTTGTAACCGCTGCGTACGCTGCGCGAGCTCAATA 10953
QY 1081 AAAATGTGA 1089
DB 10954 AAAATGTGA 10962

RESULT 2
US-09-881-457A-3
Sequence 3, Application US/09881457A
Patent No. US20020081316A1
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D
APPLICANT: Cook, Stephanie M
APPLICANT: Wild, Martha A
FILE REFERENCE: SY01105K10KOR
CURRENT APPLICATION NUMBER: US/09/881,457A
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: 09/426,352
PRIOR FILING DATE: 1999-10-25
PRIOR APPLICATION NUMBER: 08/804,372
PRIOR FILING DATE: 1997-02-21
PRIOR APPLICATION NUMBER: PCT/US95/10245
PRIOR FILING DATE: 1995-08-09
PRIOR APPLICATION NUMBER: 08/663,566
PRIOR FILING DATE: 1996-06-13
PRIOR APPLICATION NUMBER: 08/288,065
PRIOR FILING DATE: 1994-08-09
PRIOR APPLICATION NUMBER: PCT/US93/05681
PRIOR FILING DATE: 1993-06-14
PRIOR APPLICATION NUMBER: 08/023,610
PRIOR FILING DATE: 1993-02-26
PRIOR APPLICATION NUMBER: 07/898,087
PRIOR FILING DATE: 1992-06-12
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 3605
TYPE: DNA
ORGANISM: Infectious Latyngotracheitis Virus
FEATURE:
NAME/KEY: CDS
LOCATION: (585)..(1889)


```

1 LOCATION: 3694..5124
2 OTHER INFORMATION:
3 FEATURE:
4 NAME/KEY: CDS
5 LOCATION: 5210..7081
6 OTHER INFORMATION:
7 FEATURE:
8 NAME/KEY: CDS
9 LOCATION: 7245..8123
10 OTHER INFORMATION:
11 FEATURE:
12 NAME/KEY: CDS
13 LOCATION: 8333..11290
14 OTHER INFORMATION:
15 FEATURE:
16 NAME/KEY: CDS
17 LOCATION: 11098..12402
18 OTHER INFORMATION:
19 FEATURE:
20 NAME/KEY: CDS
21 LOCATION: 12510..13598
22 OTHER INFORMATION:
23 FEATURE:
24 NAME/KEY: CDS
25 LOCATION: 13792..15291
26 OTHER INFORMATION:
27 FEATURE:
28 NAME/KEY: CDS
29 LOCATION: 15298..16080
30 OTHER INFORMATION:
31 FEATURE:
32 NAME/KEY: CDS
33 LOCATION: 16129..17013
34 OTHER INFORMATION:
35 FEATURE:
36 NAME/KEY: CDS
37 LOCATION: complement (17380..18216)
38 OTHER INFORMATION:
39 US-09-994-064-59

```

Db	12931	TACGTATTGTGGCGCTAGACGACGACCAAGAAACCATTAGCTCTTCGGGATCCAACTA	129990
OY	481	TCGGTGTATCAATTCGCGAAACACCGCCGCGACTCGCGGACTTAATTCGAAGGCTTCGTGT	540
Db	12991	TCGGTGTATCAATTCGCGAAACACCGCCGCGACTCGCGGACTTAATTCGAAGGCTTCGTGT	13050
OY	541	CGCACCTTCGGATTACTACCGTCCAACTTGAGGCTATCTCAGAGACCGAGAAATGTGG	600
Db	13051	CGCACCTTCGGATTACTACCGTCCAACTTGAGGCTATCTCAGAGACCGAGAAATGTGG	13110
OY	601	CGCAACTGGCAAGCGTATGTCACAGAGGCGCACGACACAGCCGCGAGGCGAGCAACC	660
Db	13111	CGCAACTGGCAAGCGTATGTCACAGAGGCGCACGACACAGCCGCGAGGCGAGCAACC	13170
OY	661	CCGACGCCGCTCACATGCGAACAGCGCCTCCGAACTTGAAACGGAGACATTTACCTTTCCC	720
Db	13171	CCGACGCCGCTCACATGCGAACAGCGCCTCCGAACTTGAAACGGAGACATTTACCTTTCCC	13230
OY	721	TGCGTAAAAAAATGGCGCTGGAATCATTTACGAACCGACACCCGCAAAAGAAAAATTTCAACGTT	780
Db	13231	TGCGTAAAAAAATGGCGCTGGAATCATTTACGAACCGACACCCGCAAAAGAAAAATTTCAACGTT	13290
OY	781	ACTGTCCGTCCTCGGGCAATAGACGCCCTACGCTAATTGGGGTAACCTGTGCGCGTCGTG	840
Db	13291	ACTGTCCGTCCTCGGGCAATAGACGCCCTACGCTAATTGGGGTAACCTGTGCGCGTCGTG	13350
OY	841	AGCGCAACGATCGGCGCTGTCATTTGTAAATTTCCATCGTCACCGAAGAAATGTGACCCCG	900
Db	13351	AGCGCAACGATCGGCGCTGTCATTTGTAAATTTCCATCGTCACCGAAGAAATGTGACCCCG	13410
OY	901	CACCGAAATTTAGACACAGGTCCTGGCAAGACAGAGAAAGCAAGCTCCCAAACTGGAAGGGAA	960
Db	13411	CACCGAAATTTAGACACAGGTCCTGGCAAGACAGAGAAAGCAAGCTCCCAAACTGGAAGGGAA	13470
OY	961	TCGCGAAATTTGGACCCATGTTGGCTGGCGGAATTAACAAAGGCGCTGACACGATAGT	1020
Db	13471	TCGCGAAATTTGGACCCATGTTGGCTGGCGGAATTAACAAAGGCGCTGACACGATAGT	13530
OY	1021	GAACCTTGTGAACTGTTGCGATTTGTTAAACCGCTGCGCTAAGCTCGGCCGACTCAATA	1080
Db	13531	GAACCTTGTGAACTGTTGCGATTTGTTAAACCGCTGCGCTAAGCTCGGCCGACTCAATA	13590
OY	1081	AAATGTGA 1089	
Db	13591	AAATGTGA 13599	

61	ATGGGAATCGTGAATCACTGGAAATCAACAGTCTCCGCGAGATTGACAGCATCAACATCGG	120	
Db	12571	ATGGGAATCGTGAATCACTGGAAATCAACAGTCTCCGCGAGATTGACAGCATCAATCGTG	12638
QY	121	ATCGTCGCGCGCTGCGCCCGCAAGCTACAAATTCAACCTGACACTAATTTTTCATGCGTGGCAG	180
Db	12631	ATCGTCGCGCGCTGCGCCCGCAAGCTACAAATTCAACCTGACAGTAAATTTTTCATGCGTGGCAG	12690
QY	181	AGACCCGCAAAACCTTAATCTAGGAACCGTCCGCGTGGCTTTCGGTCTGATATTAACAAC	240
Db	12691	AGACCCGCAAAACCTTAATCTAGGAACCGTCCGCGTGGCTTTCGGTCTGATATTAACAAC	12750
QY	241	CAGTGGTACACGAGCACTTGAACGAGAGCGCTTGGAAATTTGCACATCATGATGCTCTCT	300
Db	12751	CAGTGGTACACGAGCACTTGAACGAGAGCGCTTGGAAATTTGCACATCATGATGCTCTCT	12810
QY	301	GTTTGTGTGCGCTGTAAGAATGACCGAGTACACGTTCTCCGCGCTCGAAGACAGTAACCCGA	360
Db	12811	GTTTGTGTGCGCTGTAAGAATGACCGAGTACACGTTCTCCGCGCTCGAAGACAGTAACCCGA	12870
QY	361	CGTCCACACCCGTTTAAAGCTCACTATACGAATCTGTCGGAACGACAGCGGATGTTTC	420
Db	12871	CGTCCACACCCGTTTAAAGCTCACTATACGAATCTGTCGGAACGACAGCGGATGTTTC	12930
QY	421	TACGTAATTTTGTGGCTAGACGACACCAAGAACCATTGAGCTCTTCCGATTCACAATA	480


```
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/468,190
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 534 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..533
US-09-994-064-18

Query Match
Best Local Similarity 31.8%; Score 346; DB 9; Length 534;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 TTACGACCGACACCGCGCAAGCAAAATTTCAACGTTACTGTCCTGCTGGGACAAATGAG 803
DB 534 TTACGACCGACACCGCGCAAGCAAAATTTCAACGTTACTGTCCTGCTGGGACAAATGAG 475
QY 804 CCTACGCTAATTTGGGGTAACCGTGGCTGGTGTGAGGCAAGCAGTGGGCTGTGAT 863
DB 474 CCTACGCTAATTTGGGGTAACCGTGGCTGGTGTGAGGCAAGCAGTGGGCTGTGAT 415
QY 864 TGTAAATTCATGCTGACCAAGCAAAATTTGACACCGCTGTC 923
DB 414 TGTAAATTCATGCTGACCAAGCAAAATTTGACACCGCTGTC 355
QY 924 GCAAGACGACGAGAAAGAGTTCCTCAAACTAGAAAGGAATCGGAAATTTGACCGCATGAT 983
DB 354 GCAAGACGACGAGAAAGAGTTCCTCAAACTAGAAAGGAATCGGAAATTTGACCGCATGAT 295
QY 984 TGGCTGCGAAATTAACAAGAGGGGCTGACGAGATAGTGAACCTGTGGAACGCTTGGCAT 1043
DB 294 TGGCTGCGAAATTAACAAGAGGGGCTGACGAGATAGTGAACCTGTGGAACGCTTGGCAT 235
QY 1044 TGTAAACCGCTGCTGCGTAAGCTGCGCCGACTCAATTAATAATGTGA 1089
DB 234 TGTAAACCGCTGCTGCGTAAGCTGCGCCGACTCAATTAATAATGTGA 189

RESULT 5
US-10-123-155-358/c
Sequence 358, Application US/10123155
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin J.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
```

```
APPLICANT: Matanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33081C30
CURRENT APPLICATION NUMBER: US/10/123,155
PRIORITY FILING DATE: 2002-04-15
PRIORITY APPLICATION REMOVED - see Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 358
LENGTH: 1049
TYPE: PRT
ORGANISM: Homo Sapien
US-10-123-155-358

Query Match
Best Local Similarity 3.3%; Score 36; DB 9; Length 1049;
Matches 35; Conservative 215; Mismatches 529; Indels 1; Gaps 1;

QY 270 CTTTGAAATTTGACTCATGATGCTCTTTGTTGCGGCTGTAAGTGAACGACGAGTA 329
DB 818 C.CBGBHAT.R..AB.ADNBMB.WETHAG..YMD...NM..NN..BN..SAM..K.. 759
QY 330 CAGCTTCCGCTCGAAGACGTAACCGACCTCCACCCGTTAAGTCACTATPACG 389
DB 758 MNSS.H.RY..TH..RMA..SY..NNM..NN.SYSGNS.Y..BAA..NDS..H.A..NM 699
QY 390 AAATCCTC-GTCCGAACGACGCGGATGTTCTACGTAATTTGCTGAGACGACACCA 448
DB 698 .G..MMMS..SM..CMTT..S.NM..N..KCH..BCS...S.SNMS..H...M..NM...RYNH 639
QY 449 AAGAACCCATTTGACCTTTCCGATCCCAATGATGATCAATTCGCAACGCGCG 508
DB 638 C.YW..BH..DNCY...AY..S..KAYSASS..HHNNK..M..BM..NMA..N..KDA..C..S..R 579
QY 509 CGACTCGCGACTATTTCCAGGCTTCGTGTGACACCTTCGATTAAGTCACTACGTCAC 568
DB 578 DSNSS..H..B..MD...TASD..H..YNNH..H..RY..T...SCN..A..S..NCS..N..GM.. 519
QY 569 TTGAGGCTATCTGACGACGAGCAAGTGGCGCAATGCGCAACGACGATGGCCACG 628
DB 518 .S..D..HSSMB...SNMS..H..A..CRMGS..NBSK..ST..NMN..YGYTHMRY..RD...B 459
QY 629 AGGCCACGACGACGACGCGCGGAGGCGCAACCGCGCGCTGCTGCAACGACGCGCT 688
DB 458 ..RS..BSATYNSG..CB..SSHCS..S..MNSB..H..BM..YM..M..KS..NT..M..MAC..H..B.. 399
QY 689 CCGAAGTTAAGCGAAGACCTTTACCTTTCCCTGCTAGAAATGCGCTGATCATTAAG 748
DB 398 N..ND..S..N..SM..M..BCT..Y..M..SM..SS..T..S..NKSTYRB...N..S..H...S... 339
QY 749 AACCGACCGCGCAAAAGCAAAATTAACGTTACTGCTGCTGCGGACAAATGAGCCCTA 808
DB 338 .D..MTHC..MT..N..S..H...MN..NM..WY..BD..SNSD..Y..BM..A..TH..TNB... 279
QY 809 CGCTAATTTGGGTAACCTGCTGCTGCTGAGCGCAAGATCGGCTGCTGTAATGTGA 868
DB 278 .SNMG..TG...TNRGY..GNCS..H...N..NN..HH...MT..KNR..R..A..AS...BA..B 219
QY 869 TTTCATGCTCACGAAACATGTGACCCCGGACGGAATTAAGACAGGTGCTGCAAG 928
DB 218 TABNNHM..S..BW..MA..N..THM..SRBRC..NYRRON..C..R...NT...A..N..MY..S..N 159
QY 929 ACAGCAAGAGAGTTCCTCAAACTAGAAAGGAATCGGAAATTTGAGACCCATGTTGCGT 988
DB 158 NT..S...S...C...NCH..R..SM..RA..CS..SY..M...YM..GKNMNSC...BCN 99
QY 989 GCGAATTAACAAGGCGCGTACAGAGATGTGTAACCTGTGACGCTGTTGCGATGTGA 1048
DB 98 GX..H..B..DH..YD..ST..S..H..DN..A..A..NAANA..CC...A..DKHAGHB..BDNM..BH..A 39
```

```
RESULT 6
; US-10-184-644-402/c
; Sequence 402, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC227
; CURRENT APPLICATION NUMBER: US/10/184,644
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 402
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-184-644-402

Query Match      3.2%; Score 35; DB 9; Length 653;
Best Local Similarity 5.0%; Pred. No. 0.2;
Matches 24; Conservative 160; Mismatches 298; Indels 0; Gaps 0;

14 TTGAACTCTGGCTCTCTCCCTGCGGAGCGCTCGACCCCTTCGCGCGATGGAAATCGTGA 73
518 YSSYS.S.SYNMMDMDYDCCSSSSSSSSBS.HHSBSSSSASMYMM.YM.M.MY 459
74 TCACGTGAATACGCTCCGCGAGATTGACGACGATCATCGTGCCTGCGCGCTC 133
458 .M.TSA..MMAM.M.M.T.M.Y...SM.KRMD.M.TH.THM.A.H.T.TMRRS.TN. 399
134 GCCCGAAGCTACATTAACGACGCTATTTTTCATGCGCTGCGAGAGACCCACAAAC 193
398 .M.....CCY.B.A.G.RMYTNYDA.GN..T...H..TMN.SCMATR.TCYTB.T.BN. 339
194 CCTACTCAGAACCGCTCCGCGCTTCGCTGCTGATTAACAACGACCTCTACCAAG 253
338 MH...TMNR.NKTHDYCBM.RHBC.MBGM.TWSTS.MMY.TST.HH.S.NMS..CYK.S 279
254 AACTTAGCAGAGCGCTTGAATAATGCATCATCGATCGTCTCTGTTTGTGCGCT 313
278 .NS.H.C.M....BBC.NB.C.SS.KBN.RS..SNSN..BSYR...SS.BC.M. 219
314 GTAAGTGACGAGTACAGCTTCTCCGCTCGACGACTAACCGACCTCCACACCGCT 373
218 CS.D...SSSR.SBT.M.DRYH.HM.CTY..HCA.RRS..HCDNSDS.BHY..G.TA... 159
374 TTAAGCTACTATAGAAATCTCGTCCGACGAGCGGATGTTCTAGTAATGTTC 433
158 DT.HRK.CSC.G..BK..C.Y.SS.YC.B.HCY..HY...YHKS...K.....K.T 99
434 GGCTAGACGACCAAGAAGCAATTCGATTCGCGATTCACATTCGATGATCAAT 493
98 RDH.SH.N..SAT.THSWMS.T...HTWMS..M...NBYM.MYC...D...CYT..T 39
OY 494 TC 495
DB 38 TS 37
```

```
RESULT 7
; US-10-184-634-402/c
; Sequence 402, Application US/10184634
```

```
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430RIC217
; CURRENT APPLICATION NUMBER: US/10/184,634
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 402
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-184-634-402

Query Match      3.2%; Score 35; DB 9; Length 653;
Best Local Similarity 5.0%; Pred. No. 0.2;
Matches 24; Conservative 160; Mismatches 298; Indels 0; Gaps 0;

14 TTGAACTCTGGCTCTCTCCCTGCGGAGCGCTCGACCCCTTCGCGCGATGGAAATCGTGA 73
518 YSSYS.S.SYNMMDMDYDCCSSSSSSSSBS.HHSBSSSSASMYMM.YM.M.MY 459
74 TCACGTGAATACGCTCCGCGAGATTGACGACGATCATCGTGCCTGCGCGCTC 133
458 .M.TSA..MMAM.M.M.T.M.Y...SM.KRMD.M.TH.THM.A.H.T.TMRRS.TN. 399
134 GCCCGAAGCTACATTAACGACGCTATTTTTCATGCGCTGCGAGAGACCCACAAAC 193
398 .M.....CCY.B.A.G.RMYTNYDA.GN..T...H..TMN.SCMATR.TCYTB.T.BN. 339
194 CCTACTCAGAACCGCTCCGCGCTTCGCTGCTGATTAACAACGACCTCTACCAAG 253
338 MH...TMNR.NKTHDYCBM.RHBC.MBGM.TWSTS.MMY.TST.HH.S.NMS..CYK.S 279
254 AACTTAGCAGAGCGCTTGAATAATGCATCATCGATCGTCTCTGTTTGTGCGCT 313
278 .NS.H.C.M....BBC.NB.C.SS.KBN.RS..SNSN..BSYR...SS.BC.M. 219
314 GTAAGTGACGAGTACAGCTTCTCCGCTCGACGACTAACCGACCTCCACACCGCT 373
218 CS.D...SSSR.SBT.M.DRYH.HM.CTY..HCA.RRS..HCDNSDS.BHY..G.TA... 159
374 TTAAGCTACTATAGAAATCTCGTCCGACGAGCGGATGTTCTAGTAATGTTC 433
158 DT.HRK.CSC.G..BK..C.Y.SS.YC.B.HCY..HY...YHKS...K.....K.T 99
434 GGCTAGACGACCAAGAAGCAATTCGATTCGCGATTCACATTCGATGATCAAT 493
98 RDH.SH.N..SAT.THSWMS.T...HTWMS..M...NBYM.MYC...D...CYT..T 39
OY 494 TC 495
DB 38 TS 37
```

```
RESULT 8
; US-10-040-739-899
; Sequence 899, Application US/10040739
; Patent No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
```

```

; MCCoy, John
; Lavallee, Edward
; Racie, Lisa
; Merberg, David
; Treacy, Maurice
; Spaulding, Ylki
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
; NUMBER OF SEQUENCES: 1519
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/040,739
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/036,520
; FILING DATE: 03-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 899:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 899:
US-10-040-739-899

Query Match 3.2%; Score 34.6; DB 9; Length 475;
Best Local Similarity 50.9%; Pred. No. 0.24; Indels 0; Gaps 0;
Matches 82; Conservative 0; Mismatches 79;

154 CTGCAGCATTTTTCAGTCTGCTGCGCAGACCCGACAAACCTTACACAGAACCTCCGC 213
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
239 CTGTACTATGTCTGCTGCGCAGAGGCTGTGTTGTCTAGAGTGTGAGTGTGCGCCAGACAA 298
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
214 GTCCGCTTTCGCTGTGATTAACAACAGTGTCTACAGAGAACTTACGAGAGCGCTTT 273
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
239 ATGGGCTGCTGCTGCTGCTCCACAAATCTTTATCCCATCTTCATCGCAAGAGGCGACT 358
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
214 GAAATTCACATCATCATGCTCTCTCTGTTTGTGCGCTG 314
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
239 GCCAAGTCTCTCAGCCAGAGTCTGTGTCTAAGGCTCTG 399
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

RESULT 9
US-09-918-995-21276/c
; Sequence 21276, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
```

```

; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21276
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(493)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-21276

Query Match 3.2%; Score 34.6; DB 9; Length 493;
Best Local Similarity 50.9%; Pred. No. 0.24; Indels 0; Gaps 0;
Matches 82; Conservative 0; Mismatches 79;

154 CTGCAGCATTTTTCAGTCTGCTGCGCAGACCCGACAAACCTTACACAGAACCTCCGC 213
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
474 CTGTACTATGTCTGCTGCGCAGAGGCTGTGTTGTCTAGAGTGTGAGTGTGCGCCAGACAA 415
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
214 GTCCGCTTTCGCTGTGATTAACAACAGTGTCTACAGAGAACTTACGAGAGCGCTTT 273
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
414 ATGGGCTGCTGCTGCTGCTCCACAAATCTTTATCCCATCTTCATCGCAAGAGGCGACT 355
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
214 GAAATTCACATCATCATGCTCTCTCTGTTTGTGCGCTG 314
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
239 GCCAAGTCTCTCAGCCAGAGTCTGTGTCTAAGGCTCTG 314
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

RESULT 10
US-10-184-644-312/c
; Sequence 312, Application US/10184644
; Publication No. US2003004930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 312
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-644-312

Query Match 3.1%; Score 34.2; DB 9; Length 802;
Best Local Similarity 6.3%; Pred. No. 0.43; Indels 0; Gaps 0;
Matches 46; Conservative 215; Mismatches 467;

350 GACTAACCGGACCTCCACACCGGTTTAACTCTATGAAATCTGTCGACAGACA 409
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
762 KMR.M.H.N.HG.B.DDBDBCB..SY.AM.CT.BBMN..HHH..CWC..SBH.CBYB. 703
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
410 GCGGAGTGTCTACGATTTGTCGGCTAGACACACCAAGAACCATGAGTCTCTG 469
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
702 HCMR.GA.C.CNRY.CAM.A.T.DNHS.BMCSRB.M.HR.S.....R..C.BANTYCY 643
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
470 CGATCCAACTATCGGTGATATCAATTCGCGAAGACGCGGACGTCGGAGCTATTCGA 529
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
642 ....AA.B.HBHG.....S.NN..SSCB..T.TYS..CSB....KTM.RM.YRHYK..B. 583
||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
```

Oy	530	AGGCTGTGTCGGACACCTTCGATTAACCTACCGTCCCAACTGGAGCCATATCGAGGACG	589
Db	582	МТМН.Н\$И\$Н. Б\$М\$Н.Н. ГАМ. НК\$К.У\$В.К\$Н.У\$С. N.ММММ. MD.\$M	523
Oy	590	AGGAAAGTTGGCGCAACTGCGCAAGCCTGATCTGGCCACGAGGAGCCACGACGACACGCGCG	649
Db	522	SC. B.ММ.М\$М\$. S.С\$С\$. N.УМТ. Т\$Н. D.В. MS.A. RTDYB.В\$К	463
Oy	650	AGGGGACACACCCCGACGCGCCCTCACTGCAACACGAGCGCCCTCGCAACTTGAACGGGACACT	709
Db	462	ММ\$М\$М\$М\$. H. R.Т\$С\$К.Н\$У\$Р\$С\$Р. M.Н. У\$У\$АМТМТ\$Н.К.К\$В.К\$В.Н\$Н.Т.	403
Oy	710	TTACCTTCCCTGGGTAGAAATGAGCGGTGATCATTTAGAACGCAACCGGACCGCAAAAGAA	769
Db	402	У.Т\$Н\$К\$У. СН\$У\$Т\$Р.РМ\$С\$А. W. Y.У. Y\$К. S. C. H.М\$D.А\$N\$R\$M	343
Oy	770	ATTCAAAAGTTACTGTCTCGCTCGTCCGGACCAATGAGCGCCCTACGCTATTGGGTGAACCGTGG	829
Db	342	SK.B. YD. A\$U\$Y. S.М\$У\$.K\$R\$D.У\$R.Р. MN\$N.D.A.T.D.М\$N\$D.Н\$У.ROM\$N.	283
Oy	830	CTGCCCTGCTGAGGCGCAACGATCGGCGCTCGTCATTTGTAATTTCCATCTCTCAACGAAACA	889
Db	282	РН.Р. Y\$К.Р\$М\$В\$G.В\$CТ.У\$У\$В\$У\$С\$В.В\$Н.Д\$А\$Р\$К.У. G\$M\$C.D.В\$К\$У.В.У\$K\$.B\$C.	223
Oy	890	TGTGACACCCCGACGAAATTTAGACAGCGTCTGCGAAGAGGAGGAGAAAGTTGCCAAA	949
Db	222	С\$К\$G.Н.С. Т.М\$К.А\$C.С\$А.С. SS\$N.SY. N. У\$Н\$М.В\$Н\$Н\$Т\$У.В.Р\$M	163
Oy	950	CTAGAAGGAATTCGAAATTTGGAGCCCATGTTGCTGCGAATTAACAAGGCGCGTG	1009
Db	162	H.Р\$H\$D\$K\$M.К\$K.С\$M\$M. R\$H\$N\$B\$У.В\$В. B\$A\$N\$C.S\$S. B\$M\$C.А\$M\$M\$A\$P\$T\$B	103
Oy	1010	ACCAAGATAGGAACCTTGCGAACTGCTGGCGATTTGAACCGCTCTGCGSTAAGCTGCC	1069
Db	102	Т\$У\$А. Р\$M.Т\$А\$В\$С\$В. NY\$H.С\$H.Н\$C. W. Т\$У\$Н\$У\$C\$T\$C\$P.СТ\$T.С\$У\$C\$.ST	43
Oy	1070	CGGACTCA	1077
Db	42	У\$У.С\$Т\$М	35

```

RESULT 11
US-10-184-634-312/c
: Sequence 312, Application US/10184634
: Publication No. US2003006864A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Collin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zhenli
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: TITLE OF INVENTION: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3430R1C217
: CURRENT APPLICATION NUMBER: US/10/184,634
: CURRENT FILING DATE: 2002-06-28
: PRIOR APPLICATION removed - See file Wrapper or Palm
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 312
: LENGTH: 802
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-184-634-312

```

```
Query Match      3.1%; Score 34.2; DB 9; Length 802;
Best Local Similarity 6.3%; Pred. No. 0.43;
```

	Matches	46/	Conservative	215/	Mismatches	467/	Indels	0/	Gaps	0/
OY	350	GACTAACCGGACCTCCACACCCCGTTTAAGCTCATATACGAAATCTGTGCGCAACGACA	409							
Db	762	KMR..M..H..N..HG..B..DDBDBC..SV..AM..CT..BBMN..HBH..CWC..SBH..CCVB..B	703							
OY	410	GGCGGATGTCTAGCTAATATTGTTCGGGTAGACGACACCAAGAACCCTTGTAGCTGTTCG	469							
Db	702	HCMR..GA..C..CRNTW..CAM..A..T..DNHS..BMCSEB..M..HR..S.....R...C..BANTGY	643							
OY	470	CGATCCAACTATTCGGGTATCAATTCGCGGAACACCGCCGCACTGGCGGACTATTTCCA	529							
Db	642	...AA..B..HBHG...S..NN...SSCB...T..TYS...CSB...KHTM..RM..YRHYK...B..	583							
OY	530	AGGCTGCTGTTCGACACTTCGTGGATTACTACCGTCCCACTTGAGGCGCATCTCAGACCG	589							
Db	582	MTMH..NHSN...BB..MBN..N...CAM...NCK..YB..KH..YCS...N..MMHM..MD..SM	523							
OY	590	AGGAAGTTGGCGCACTGCGCAAGCGTACTGTGCGCACGAGAGCCACGACGACGACCGCG	649							
Db	522	SC.....B..MM..SNS...S..SCS...N..YMT...TH...D..B...MS..A...RTDYB..BAK	463							
OY	650	AGCGGACACACCCCGACGCCCTCACTGCAACCCAGCGCCTCCGAACCTGAGCGGAACCT	709							
Db	462	MMCDMMR...H..R..TCRK..NBYRCR...M..H..YCYAMTNTN..K..KB..KH..H..T..	403							
OY	710	TTACTCTCCCTGGCTAGATAATGCGCGGTGATCATTTAGAACCGACCCGCAAAACGAA	769							
Db	402	Y..TSHKCY...CHBTSR..RMCA...W.....Y..Y...YK..S...C...H..MD..ANSRM	343							
OY	770	ATTCAAAAGTACTGTCTGCTCGTCCGAGACATGAGCCCTACGCTAATTTGGGTAAACCTGTG	829							
Db	342	SK..B...YD...AYDY..S..WYS..KRSD..YR..R...NMN..D..A..T..D..MSND..HY..RCMNM	283							
OY	830	CTGCCGTCGTAGCGGCAACGATCGGCCCTCGTCATTTGTAATTTCCATCTGTCACACGAAACA	889							
Db	282	..RN..R...YK..RSMWG..BCT..YUBGYCB..BH..DARK..Y..GWC..D..BKYU..B..YSK..BC	223							
OY	890	TGTGACCCCGCACCGAATTTAGACACGAGTCTGCAAGACGACGACGAAGACGTTGCCAAA	949							
Db	222	CCRG..N..C...T..MCK..AAC..CA..C.....SSN..SY...N...YHOM..RBHNTYK..W..RM	163							
OY	950	CTAGAAGGAATCGGAAATTTGGACCCATGTGCTGCGAATAAACAAGGCGCGTG	1005							
Db	162	H..RHDKRM..KK..SMM...RSHHBY...BB...BSAHS..SS...BMC..AMSNAYRRTTB	103							
OY	1010	ACCAAGATAGCAACTGTGCACTGGTTCGATTTGTAACCCGCTCGCTAGCTAGCTGCG	1066							
Db	102	TYA...RM..TAKCB...NYHY..CCH..HSC...W...TCYHCTYCCP..STTT..CSYGC...ST	43							
OY	1070	CCGACTCA	1077							
Db	42	YY..CTYM	35							

```

RESULT 12
US-10-125-540-594
: Sequence 594, Application US/10125540
: Publication No. US2003005875A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: P12141
: CURRENT APPLICATION NUMBER: US/10/125,540
: CURRENT FILING DATE: 2002-04-19
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 646
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 594
: LENGTH: 2855
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-125-540-594

```

```
Query Match 3.1%; Score 34; DB 9; Length 2855;
Best Local Similarity 57.5%; Pred. No. 0.98;
Matches 61; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 596 GTTGGCGCACTGCAAGCTTACGTTGGCCAGGAGGCGCCAGCAGCGCGGAGCGCA 655
DB 2122 GGTGGCCAGCGGCGGAGCGGCGGCGGCGGAGGCGCGGAGCGCGGAGCGCGGCT 2181

QY 656 CAACCCGAGCGCGCTCAGTCAACCGAGCGCGCTCGAAGTGAAGC 701
DB 2182 CCCACCGGCTCCCTCCAGCTGAGAGCATGATGACGAGAGC 2227

RESULT 13
US-09-764-870-594
; Sequence 594, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PCT/14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION DATA REMOVED - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 594
; LENGTH: 2855
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-870-594

Query Match 3.1%; Score 34; DB 10; Length 2855;
Best Local Similarity 57.5%; Pred. No. 0.98;
Matches 61; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 596 GTTGGCGCACTGCAAGCTTACGTTGGCCAGGAGGCGCCAGCAGCGCGGAGCGCA 655
DB 2122 GGTGGCCAGCGGCGGAGCGGCGGCGGCGGAGGCGCGGAGCGCGGAGCGCGGCT 2181

QY 656 CAACCCGAGCGCGCTCAGTCAACCGAGCGCGCTCGAAGTGAAGC 701
DB 2182 CCCACCGGCTCCCTCCAGCTGAGAGCATGATGACGAGAGC 2227

RESULT 14
US-10-022-832-1
; Sequence 1, Application US/10022832
; Patent No. US20030039896A1
; GENERAL INFORMATION:
; APPLICANT: COUTURE, JOSE
; APPLICANT: HAMEL, JOSE
; APPLICANT: BRODEUR, BERNARD R.
; TITLE OF INVENTION: NOVEL CHLAMYDIA ANTIGENS AND CORRESPONDING DNA
; FILE REFERENCE: BIOVAC-15
; CURRENT APPLICATION NUMBER: US/10/022,832
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/256,941
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-10-022-832-1

Query Match 3.0%; Score 33.2; DB 9; Length 777;
Best Local Similarity 56.4%; Pred. No. 0.94;
Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
```

```
QY 929 ACAGCAAGACGTTCCCAAGTAGAAGGAAATGCGAATTTGGACCCATGGTGGCT 988
DB 134 ATGACATTAATAATATCCGATATATCAAGATTAATGCAAGCAATTTGCTATTATGAC 193

QY 989 GCGAATTAACAAGGCGCTCAGCAGATATGATGACTTGTGAGACTGTT 1038
DB 194 GCGAATTAACGCAAGTACAGATATGTTTGTATATTCAGAAAGTGGCT 243

RESULT 15
US-09-864-761-6494/c
; Sequence 6494, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecm1ga-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 6494
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004889.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
```

